UNIVERSE

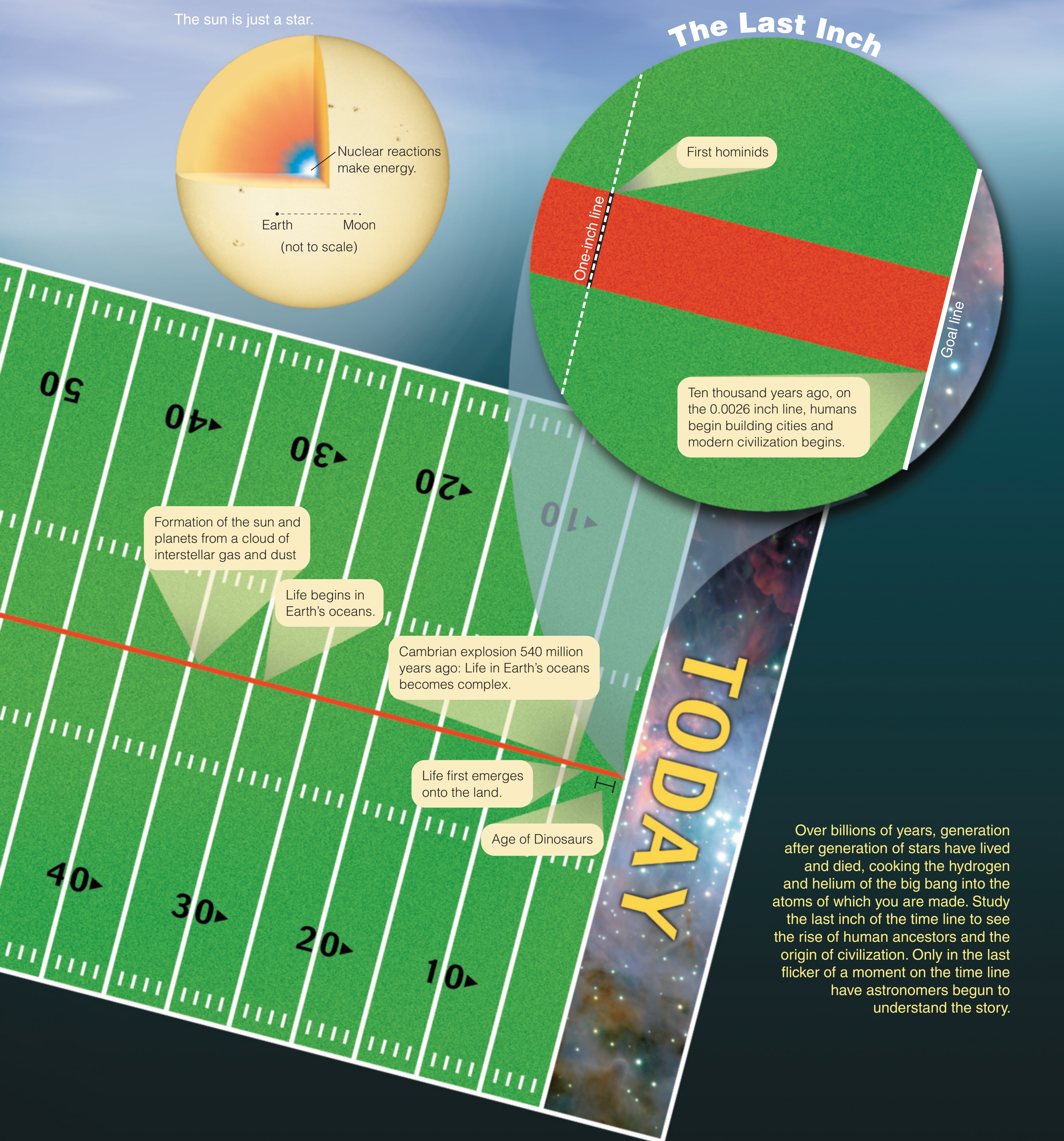
8TH EDITION

SOLAR SYSTEMS, STARS AND GALAXIES

MICHAEL SEEDS AND DANA BACKMAN

Imagine the history of the universe as a time line down the middle of a football field. The story begins on one goal line as the big bang fills the universe with energy and a fantastically hot gas of hydrogen and helium. Follow the history from the first inch of the time line as the expansion of the universe cools the gas and it begins to form galaxies and stars.

A typical galaxy contains 100 billion stars.



Over billions of years, generation after generation of stars have lived and died, cooking the hydrogen and helium of the big bang into the atoms of which you are made. Study the last inch of the time line to see the rise of human ancestors and the origin of civilization. Only in the last flicker of a moment on the time line have astronomers begun to understand the story.

© Cengage Learning 2014; images: Anglo-Austrialian Observatory/David Malin images

ENHANCED WebAssign

Increased Engagement. Improved Outcomes. Superior Service.

Exclusively from Cengage Learning, **Enhanced WebAssign®** combines the exceptional Physics and Astronomy content that you know and love with the most powerful and flexible online homework solution, WebAssign. **Enhanced WebAssign** engages students with immediate feedback, rich tutorial content, and interactive eBooks—helping students to develop a deeper conceptual understanding of their subject matter. Online assignments can be built by selecting from thousands of text-specific problems or supplemented with problems from any Cengage Learning textbook in our collection.

With Enhanced WebAssign, you can

- Help students stay on task with the class by requiring regularly scheduled assignments using problems from the textbook.
- Provide students with access to a personal study plan to help them identify areas of weakness and offer remediation.
- Focus on teaching your course and not on grading assignments. Use **Enhanced WebAssign** item analysis to easily identify the problems that students are struggling with.

For more information visit:

www.cengage.com/ewa

CENGAGENOW

CengageNOW offers all of your teaching and learning resources in one intuitive program organized around the essential activities you perform for class—lecturing, creating assignments, grading, quizzing, and tracking student progress and performance.

- **More Control in Less Time**—Flexible and automatically graded assignments, quizzes, and Gradebook options to best fit overall course goals.
- **Delivers Better Student Outcomes**—Includes Personalized Study, a diagnostic tool (featuring chapter specific Pre-test, Study Plan, and Post-test) that empowers students to master concepts, prepare for exams, and be more involved in class.

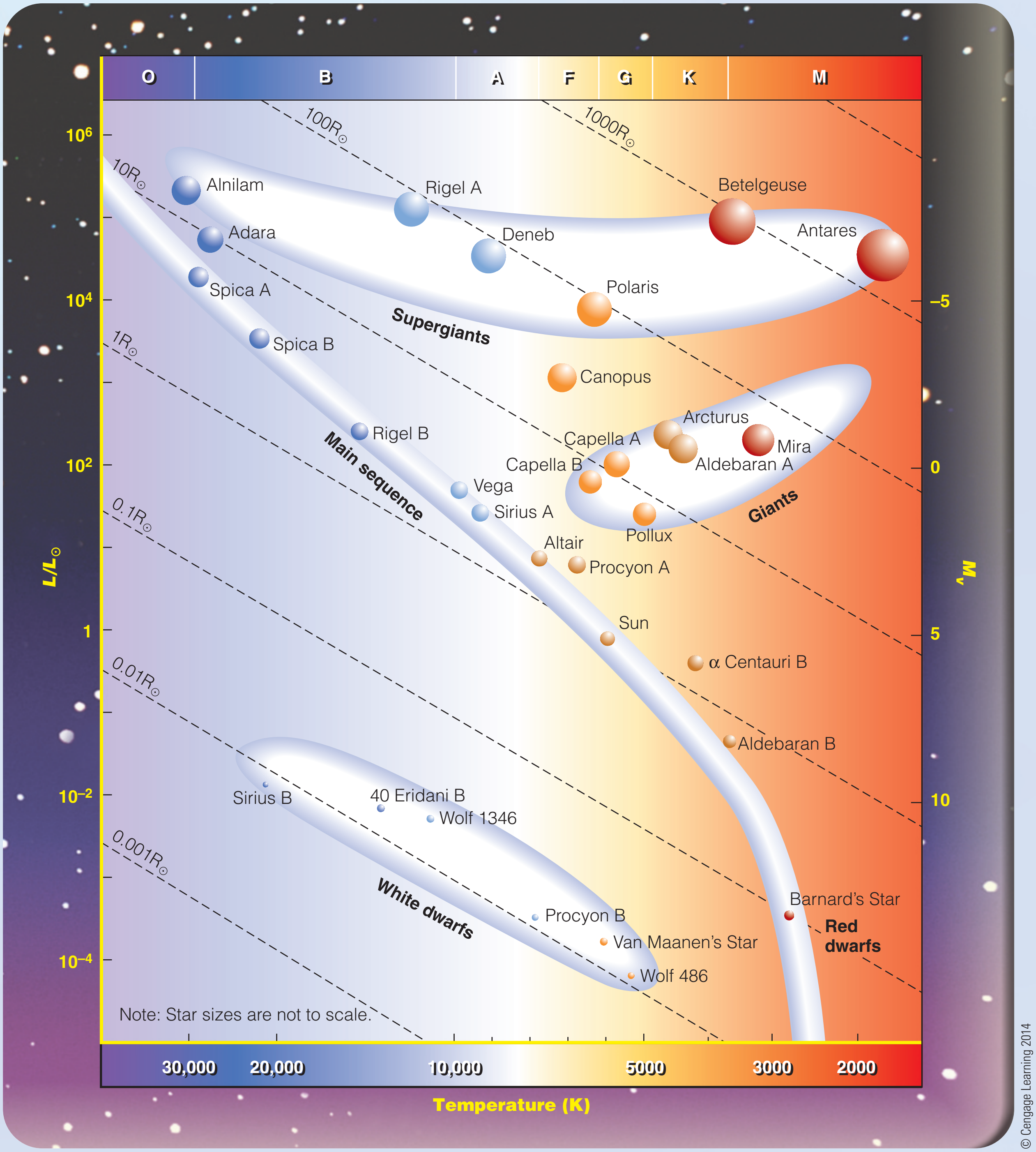
For more information:

Contact your Cengage Learning representative

TRANSFORMING LEARNING TRANSFORMING LIVES

CENGAGE Learning®

Flash Reference: H–R Diagram



The H–R diagram is the key to understanding stars, their birth, their long lives, and their eventual deaths. Luminosity ($L/L_\odot$) refers to the total amount of energy that a star emits in terms of the sun's luminosity, and the temperature refers to the temperature of its surface. Together, the temperature and luminosity of a star locate it on the H–R diagram and tell astronomers its radius, its family relationships with other stars, and a great deal about its history and fate.

Flash Reference: Comparative Planetology

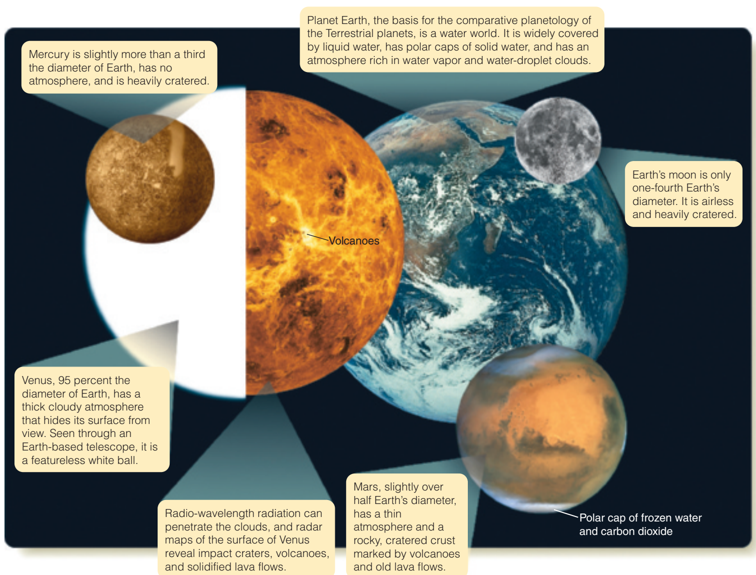
The terrestrial or Earthlike planets lie very close to the sun, and their orbits are hardly visible in a diagram that includes the outer planets.

Mercury, Venus, Earth and its moon, and Mars are small worlds made of rock and metal with little or no atmospheric gases.

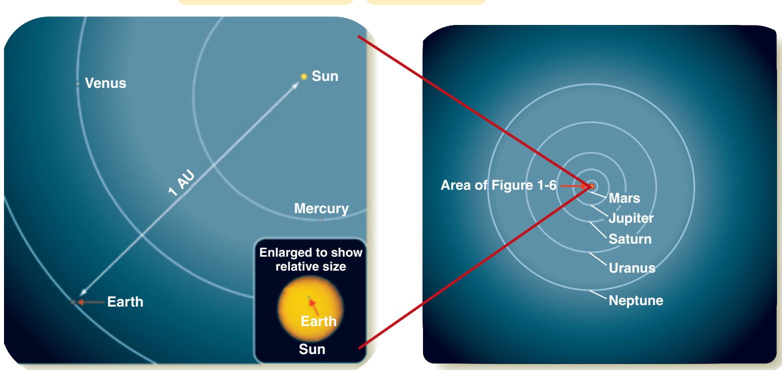
The outer worlds of our solar system orbit far from the sun. Jupiter, Saturn, Uranus, and Neptune are Jovian or Jupiter-like planets much bigger than Earth. They contain large amounts of low-density gases.

Pluto is one of a number of small, icy worlds orbiting beyond Neptune. Astronomers have concluded that Pluto is not really a planet and now refer to it as a dwarf planet.

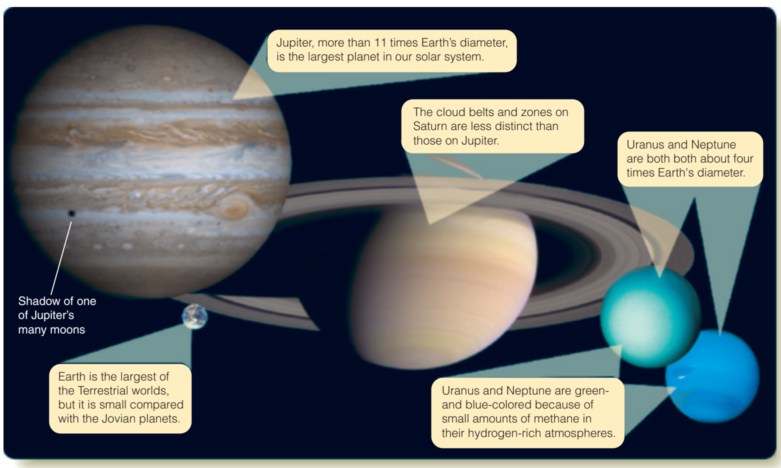
The Terrestrial Worlds



Planetary Orbits



The Outer Worlds

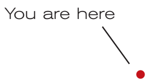


• *Horizons* readers: see page 341 for the terrestrial planets. See pages 3 and 4 for the two orbital diagrams. See page 404 for the outer worlds. *Universe* readers: See page 167 for the terrestrial planets. See pages 3 and 4 for the two orbital diagrams. See page 210 for the outer worlds.

Flash Reference: Arrows

This book is designed to use arrows to alert you to important concepts in diagrams and graphs. Some arrows point things out, but others represent motion, force, or even the flow of light. Look at arrows in the book carefully and use this Flash Reference card to catch all of the arrow clues.

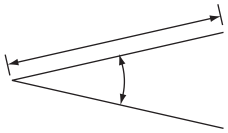
Point at things:



Force:

Process flow:

Measurement:

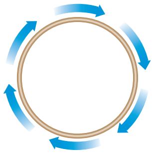


Direction:

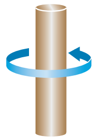
Radio waves, infrared, photons:

Motion:

Rotation 2-D

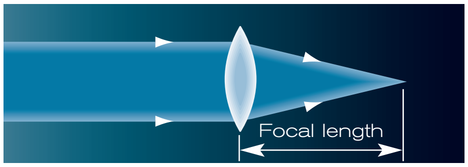


Rotation 3-D



Linear

Light flow:
Updated arrow style



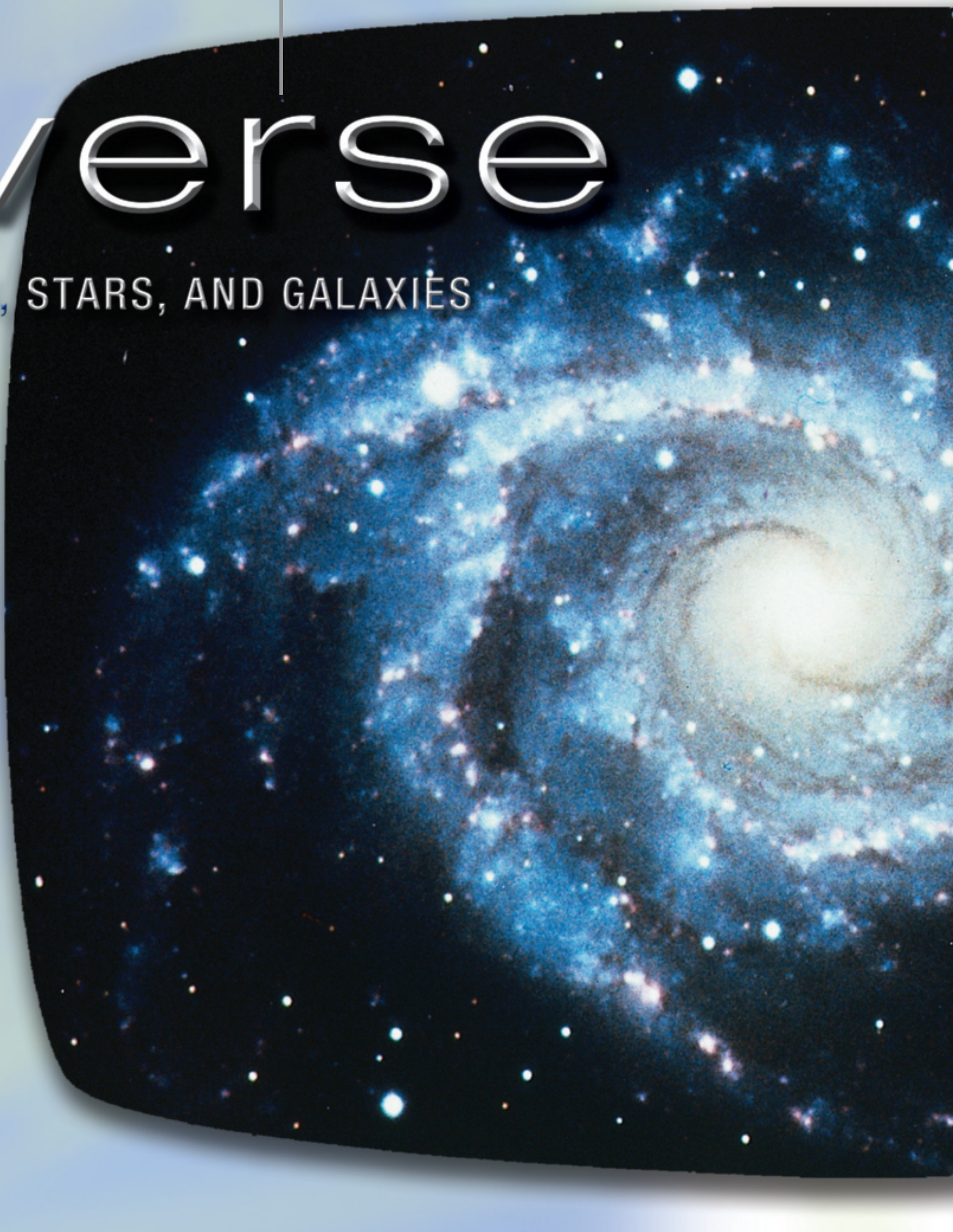
Line art on this page © Cengage Learning 2014

8

EIGHTH EDITION

Universe

SOLAR SYSTEMS, STARS, AND GALAXIES



About the Authors

Mike Seeds has been a Professor of Physics and Astronomy at Franklin and Marshall College in Lancaster, Pennsylvania, since 1970. In 1989 he received F&M College's Lindback Award for Distinguished Teaching. Mike's love for the history of astronomy led him to create upper-level courses on archaeoastronomy and changing concepts of the universe. His research interests focus on variable stars and the automation of astronomical telescopes. Mike is coauthor with Dana Backman of *Foundations of Astronomy,* Twelfth Edition (2013); *Stars and Galaxies,* Eighth Edition (2013); *The Solar System,* Eighth Edition (2013); and *ASTRO* (2010), all published by Cengage Learning. He was Senior Consultant for creation of the 20-episode telecourse accompanying the book *Horizons: Exploring the Universe.*

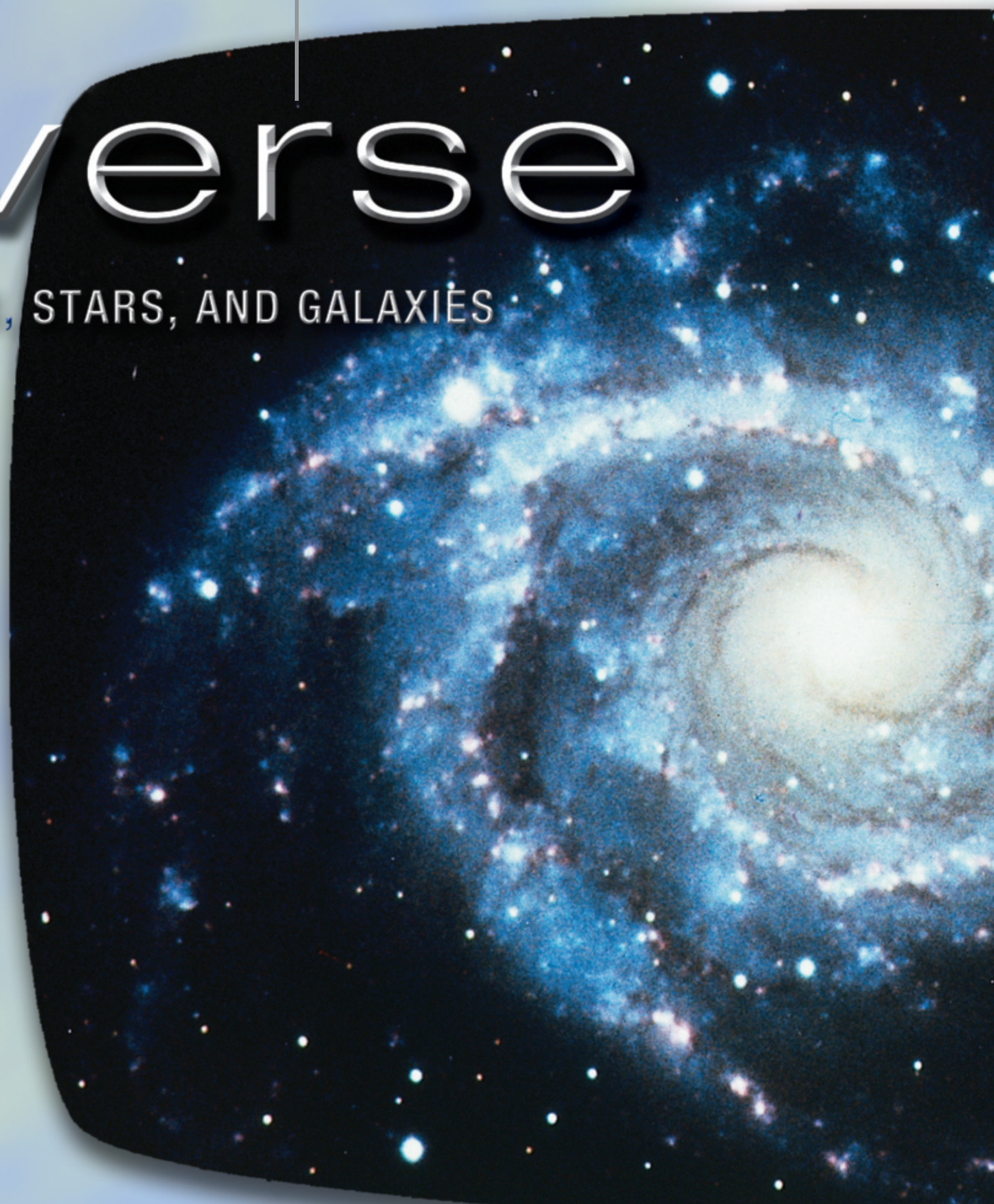
Dana Backman taught in the physics and astronomy department at Franklin and Marshall College in Lancaster, Pennsylvania, from 1991 until 2003. He invented and taught a course titled "Life in the Universe" in F&M's interdisciplinary Foundations program. Dana now teaches introductory astronomy at Santa Clara University, a course on global climate change, and a course on cosmology in Stanford University's Continuing Studies Program. Dana is employed by the SETI Institute in Mountain View, California, as the manager of Outreach (education, public outreach, and media relations) for NASA's Stratospheric Observatory for Infrared Astronomy (SOFIA) at NASA's Ames Research Center. Dana is coauthor with Mike Seeds of *Foundations of Astronomy,* Twelfth Edition (2013); *Stars and Galaxies,* Eighth Edition (2013); *The Solar System,* Eighth Edition (2013); and *ASTRO* (2010), all published by Cengage Learning.

EIGHTH EDITION

8

Universe

SOLAR SYSTEMS, STARS, AND GALAXIES



Michael Seeds

Joseph R. Grundy Observatory
Franklin and Marshall College

Dana Backman

SOFIA (Stratospheric Observatory
for Infrared Astronomy)
SETI Institute & NASA Ames
Research Center

BROOKS/COLE
CENGAGE Learning

Australia • Brazil • Japan • Korea • Mexico • Singapore • Spain • United Kingdom • United States

BROOKS/COLE
CENGAGE Learning

***Universe: Solar Systems, Stars, and Galaxies,* 8th Edition**

Michael Seeds and Dana Backman

Publisher for Science: Mary Finch

Publisher for Astronomy: Charles Hartford

Development Editor: Margaret Pinette

Associate Development Editor: Brandi Kirksey

Editorial Assistant: Brendan Killion

Media Editor: Stephanie Van Camp

Content Project Manager: Cathy Brooks

Art Director: Cate Barr

Manufacturing Planner: Sandee Milewski

Rights Acquisition Specialist: Shalice Shah-Caldwell

Production Service and Compositor: Integra Software Services

Cover Designer: Lisa Kuhn, Curio Press

Cover Image: NASA, S. Gezari (The Johns Hopkins University), and J. Guillochon (University of California, Santa Cruz)

Library of Congress Control Number: 2012950166

ISBN-13: 978-1-133-94050-0

ISBN-10: 1-133-94050-1

Brooks/Cole
20 Channel Center Street
Boston, MA 02210
USA

Printed in the United States of America
1 2 3 4 5 6 7 16 15 14 13 12

Dedication

For our families

Brief Contents

Contents

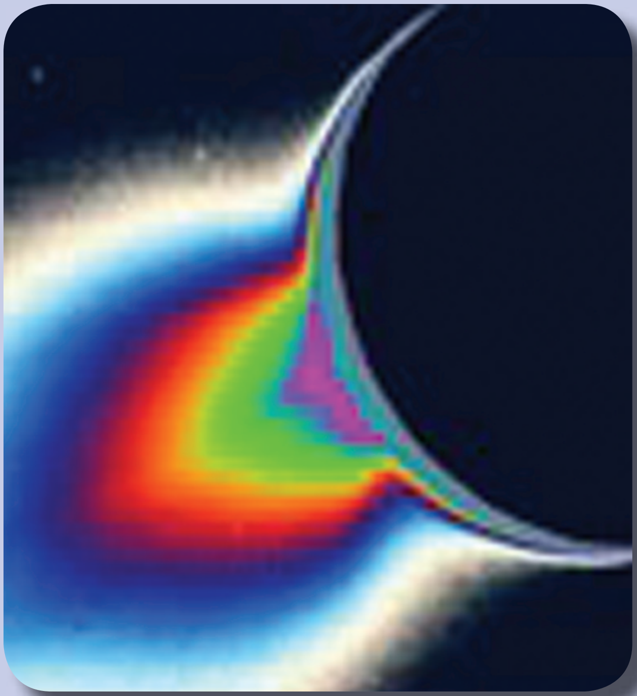
Part 1: The Sky

Part 2: The Solar System

How Do We Know?

Concept Art Portfolios

Reasoning with Numbers



Part 3: The Stars

How Do We Know?

Concept Art Portfolios

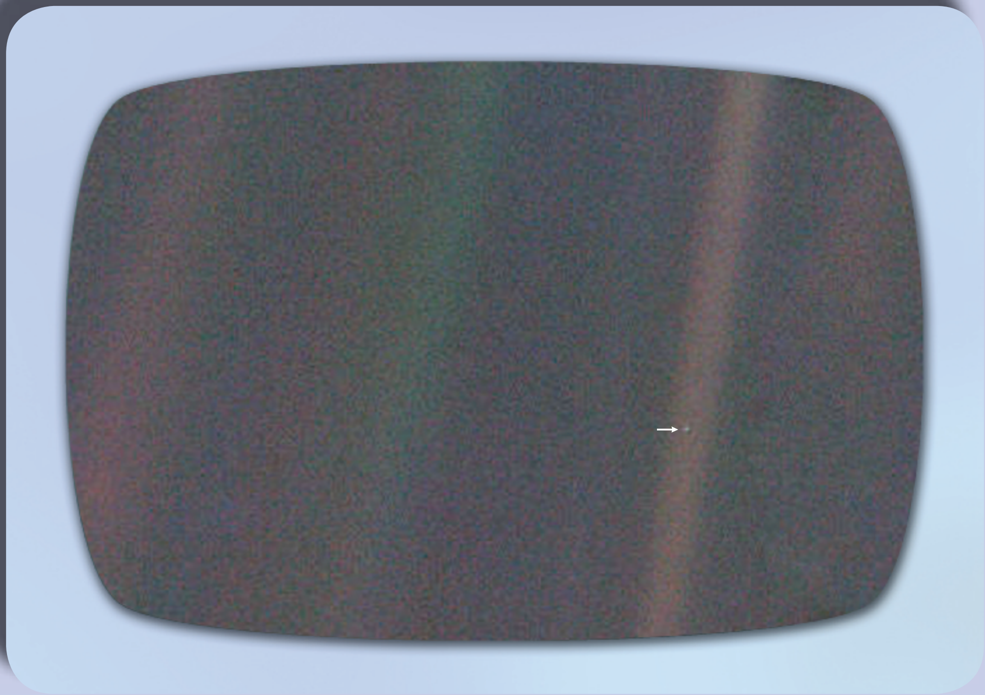
Reasoning with Numbers

Part 4: The Universe of Galaxies

Part 5: Life

How Do We Know?

Concept Art Portfolios



A Note to the Student

From Mike and Dana

We are excited that you are taking an astronomy course and using our book. You are going to see some amazing things, from the icy rings of Saturn to monster black holes. We are proud to be your guides as you explore.

We have developed this book to help you expand your knowledge of astronomy, from the moon and a few stars in the evening sky to a deeper understanding of the extent, power, and diversity of the universe. You will meet worlds where it rains methane, stars so dense atoms cannot exist, colliding galaxies that are ripping each other apart, and a universe that is expanding faster and faster.

Two Goals

This book is designed to help you answer two important questions:

- **What are we?**
- **How do we know?**

By the question "What are we?" we mean, "How do we fit into the universe and its history?" The atoms you are made of had their first birthday in the big bang when the universe began, but those atoms were cooked and remade inside stars, and now they are inside you. Where will they be in a billion years? Astronomy is the only course on campus that can tell you that story, and it is a story that everyone should know.

By the question "How do we know?" we mean, "How does science work?" What is the evidence, and how do you know it is true? For instance, how can anyone know there was a big bang? In today's world, you need to think carefully about the things so-called experts say. You should demand explanations. Scientists have a special way of knowing based on evidence that makes scientific knowledge much more powerful than just opinion, policy, marketing, or public relations. It is the human race's best understanding of nature. To understand the world around you, you need to understand how science works. Throughout this book, you will find boxes called ***How Do We Know?*** They will help you understand how scientists use the methods of science to know what the universe is like.

Expect to Be Astonished

One reason astronomy is exciting is that astronomers discover new things every day. Astronomers expect to be astonished. You can share in the excitement because we have worked hard to include the newest images, the newest discoveries, and the newest insights that will take you, in an introductory course, to the frontier of human knowledge. Huge telescopes in space and on remote mountaintops provide a daily dose of excitement that goes far beyond sensationalism. These new discoveries in astronomy are exciting because they are about us. They tell us more and more about what we are.

As you read this book, notice that it is not organized as lists of facts for you to memorize. That could make even astronomy boring. Rather, this book is organized to show you how scientists use evidence and theory to create logical arguments that show how nature works. Look at the list of special features that follows this note. Those features were carefully designed to help you understand astronomy as evidence and theory. Once you see science as logical arguments, you hold the key to the universe.

Do Not Be Humble

As teachers, our quest is simple. We want you to understand your place in the universe—not just your location in space but your location in the unfolding history of the physical universe. Not only do we want you to know where you are and what you are in the universe, but we want you to understand how scientists know. By the end of this book, we want you to know that the universe is very big but that it is described by a small set of rules and that we humans have found a way to figure out the rules—through a method called science.

To appreciate your role in this beautiful universe, you must learn more than just the facts of astronomy. You must understand what we are and how we know. Every page of this book reflects that ideal.

Mike Seeds
mseeds@fandm.edu

Dana Backman
dbackman@sofia.usra.edu

Key Content and Pedagogical Changes to the Eighth Edition

- Every chapter has been reviewed and updated with the latest discoveries and images such as photos of colliding galaxies and planets orbiting distant stars. You will read about methane lakes on Saturn's moon Titan and the newest understanding of bursts of gamma rays detected coming from the most distant galaxies. Every chapter has been reviewed and updated with the latest discoveries and images. You will read about particles from distant supernovae flying through Earth, the coevolution of galaxies with supermassive black holes, the discovery of Earth-size and Earth-temperature extrasolar planets, and findings from the first robot probes to orbit Mercury and the asteroid Vesta.
- Normal galaxies and active galaxies have been unified in a single new Chapter 18 to better show how active galactic nuclei are a natural stage in the evolution of normal galaxies. Chapter 15 ("The Deaths of Stars") has been updated with new photos and a new graph regarding the 1987A neutrino burst. Three new subsections, *Classifying Supernovae, A History of Supernovae*, and *Supernova Remnants* update and reorganize the discussion of supernovae.
- The discussion of Cyg-X-1 in Chapter 16 ("Neutron Stars and Black Holes") has been reorganized and updated to emphasize its place in the history of the subject.
- Chapter 10 uses comparative planetology to analyze the structure and history of Mercury, Venus, and Mars. Chapter 18 ("Galaxies: Normal and Active") has been updated with new images, and a new section *The Coevolution of Galaxies and Black Holes* has been added to include the newest understanding.
- The discussion of stellar spectra and their classification has been moved to Chapter 13 to better illustrate how astronomers know what stars are like. Chapter 8 ("The Origin of the Solar System") has been updated with the newest information regarding the wide and wonderful variety of extrasolar planets discovered by the *Kepler* and *Corot* space telescopes and ground-based research programs.
- Chapter 10 ("Mercury, Venus, and Mars") and Chapter 12 ("Meteorites, Asteroids, and Comets") have been updated with new findings and images from the *MESSENGER* and *Dawn* space missions, respectively.

Special Features

- ***What Are We?*** essays are placed at the end of each chapter to help you understand your own role in the astronomy you have just learned.
- ***How Do We Know?*** commentaries appear in every chapter and will help you see how science works. They point out where scientists use statistical evidence, why they think with analogies, and how they build confidence in theories.
- ***Special two-page art spreads*** provide an opportunity for you to create your own understanding and share in the satisfaction that scientists feel as they uncover the secrets of nature.
- ***Guided discovery figures*** illustrate important ideas visually and guide you to understand relationships and contrasts interactively.
- ***Guideposts*** on the opening page of each chapter help you see the organization of the book by focusing on a small number of questions to be answered as you read the chapter.
- ***Scientific Arguments*** at the end of many text sections are carefully designed questions to help you review and synthesize concepts from the section. A short answer follows to show how scientists construct scientific arguments from observations, evidence, theories, and natural laws that lead to a conclusion. A further question then gives you a chance to construct your own scientific argument on a related issue.
- ***End-of-Chapter Review Questions*** are designed to help you review and test your understanding of the material.
- ***End-of-Chapter Discussion Questions*** go beyond the text and invite you to think critically and creatively about scientific questions. You can think about these questions yourself or discuss them in class.
- ***Virtual Astronomy Labs.*** Enhance students' understanding of the scientific method with **Virtual Astronomy Laboratories** available at CengageNOW. Focusing on 20 of the most important concepts in astronomy, these labs offer students hands-on exercises that complement text topics. Instructors can set up classes online and view student results, or students can print their reports for submission, making the labs ideal for homework assignments, lab exercises, and extra-credit work.
- ***CengageNOW™*** offers all of your teaching and learning resources in one place to help teach an astronomy course. CengageNOW satisfies students who prefer to use digital resources to study. Students have access to an integrated eBook, animations, interactive quizzes, videos, and other multimedia tools to help them get the most out of their course. CengageNOW can not only help improve student performance, but also provide information you can use. Students can master key concepts and prepare for exams with CengageNOW's Personalized Study Plan—a diagnostic program plus study plan—preloaded with an integrated eBook and other text-specific material. CengageNOW also provides the end-of-chapter homework questions so you can assign and manage homework online.

- ***Enhanced WebAssign*** provides students with Web-based access to a digital version of *Universe*, Eighth Edition, and the end-of-chapter homework questions so you can assign and manage homework online. WebAssign has a customizable and interactive eBook, the **Cengage YouBook,** that lets you tailor the textbook to fit your course and connect with your students. You can remove and rearrange chapters in the table of contents and tailor assigned readings that match your syllabus exactly. Powerful editing tools let you change as much as you'd like, or leave it just like it is. You can highlight key passages or add sticky notes to pages to comment on a concept in the reading, and then share any of these individual notes and highlights with your students, or keep them personal. You can also edit narrative content in the textbook by adding a text box or striking out text. With a handy link tool, you can drop in an icon at any point in the eBook that lets you link to your own lecture notes, audio summaries, video lectures, or other files on a personal website or anywhere on the web. A simple YouTube widget lets you easily find and embed videos from YouTube directly into eBook pages. Students can also highlight the text, add their own notes, and bookmark the text. Animations and videos play right on the page at the point of learning so that they are not speed bumps to reading, but true enhancements.

New to *Universe* are **animation tutorials** written exclusively for *Universe* by text author Mike Seeds. These tutorials build on the interactive animations from the Cengage YouBook and are assignable in WebAssign. These tutorials will help students review important concepts and explore topics from the textbook in more detail. Each tutorial requires a student to consider, and sometimes manipulate an animation, and to then answer a series of questions. Hints are offered with each step, which encourage students to think through each question. Animation tutorials will build student reasoning so they will ultimately be able to draw conclusions.

Acknowledgments

Over the years we have had the guidance of a great many people who care about astronomy and teaching. We would like to thank all of the students and teachers who have contributed their insights and suggestions. Their comments have been very helpful in shaping this book.

We would especially like to thank the reviewers, whose careful analysis and thoughtful suggestions have been invaluable in completing this new edition.

Many observatories, research institutes, laboratories, and individual astronomers have supplied figures and diagrams for this edition. They are listed in the credits, and we would like to thank them specifically for their generosity.

Special thanks go to Jamie Backman, who has been an insightful reader contributing many insights to the way astronomy is organized and presented.

We are happy to acknowledge the use of images and data from a number of important programs. In preparing materials for this book, we used NASA's Sky View facility located at NASA Goddard Space Flight Center, and the SIMBAD database operated at CDS, Strasbourg, France. We have also made use of atlas images and mosaics obtained as part of the Two Micron All Sky Survey (2MASS), a joint project of the University of Massachusetts and the Infrared Processing and Analysis Center/California Institute of Technology, funded by the National Aeronautics and Space Administration and the National Science Foundation. A number of solar images are used by the courtesy of the *SOHO* consortium, a project of international cooperation between ESA and NASA. We would like to thank Michele Montgomery of The University of Central Florida for contributing the Great Debates material, as well as for her helpful comments and advice.

It is always a pleasure to work with the Cengage Learning team. Special thanks go to all of the people who have contributed to this project, including Editorial Assistant Brendan Killion, Associate Development Editor Brandi Kirksey, Media Editor Stephanie Van Camp, and Senior Content Project Manager Cathy Brooks. We have enjoyed working with Margaret Pinette, and we appreciate her understanding and goodwill. We would especially like to thank publishers Mary Finch and Charlie Hartford for their help and guidance throughout this project.

Most of all, we would like to thank our families for putting up with "the books." They know all too well that textbooks are made of time.

Mike Seeds
Dana Backman

Recent Reviewers

Mitzi Adams, NASA Marshall Center
Scott Hildreth, Chabot College
Andrea N. Lommen, Franklin and Marshall College
Chris McKay, NASA Ames
Scott Miller, Pennsylvania State University
Luisa Rebull, California Institute of Technology
Mark Reid, Harvard-Smithsonian Center for Astrophysics
Ata Sarajedini, University of Florida
Larry C. Sessions, Metropolitan Sate College of Denver

1

Here and Now

Guidepost

As you study astronomy, you also learn about yourself. You are a planet-walker, and this chapter offers you a preview of what that means. The planet you live on whirls around a star that moves through a universe filled with other stars and galaxies. You owe it to yourself to know where you are located in the universe and when you are living its history because those are important steps to knowing what you are.

In this chapter, you will consider three important questions about astronomy:

- **Where are you and Earth in the universe?**
- **How does the time span of human civilization compare with the age of the universe?**
- **Why study astronomy?**

This chapter is a summary of your upcoming exploration of deep space and deep time. The following chapter continues your journey by looking at the night sky as seen from Earth. Later chapters will provide many examples of how science gives you a way to know and understand nature.

A cloud of gas and dust in space named Messier 78, so far away its light takes 1600 years to reach Earth. This image was created by a private citizen, Igor Chekalin of Russia, in response to a worldwide contest sponsored by the European Southern Observatory (ESO) in 2010. Mr. Chekalin's prize was an all-expense-paid trip to Chile to work with the astronomers using one of ESO's giant telescopes.

Igor Chekalin

Visual-wavelength image

The longest journey begins with a single step.

—LAO-TZU

1-1 Where Are We?

To FIND YOUR PLACE among the stars, you can take a cosmic zoom, a ride out through the universe to preview the kinds of objects you are about to study.

You can begin with something familiar. ■ Figure 1-1 shows a region about 50 feet across occupied by a human being, a sidewalk, and a few trees—all objects whose size you can understand. Each successive picture in this cosmic zoom will show you a region of the universe that is 100 times wider than the preceding picture. That is, each step will widen your **field of view,** the region you can see in the image, by a factor of 100.

■ **Figure 1-2**

This box ■ represents the relative size of the previous frame.

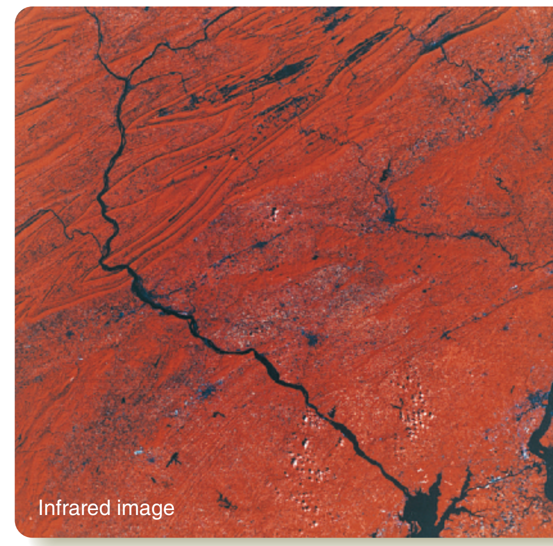
© USGS

■ **Figure 1-1**

Michael A. Seeds

Widening your field of view by a factor of 100 allows you to see an area 1 mile in diameter (■ Figure 1-2). People, trees, and sidewalks have become too small to see, but now you see a college campus and surrounding streets and houses. The dimensions of houses and streets are familiar. This is still the world you know.

■ **Figure 1-3**



NASA

Before leaving this familiar territory, you should make a change in the units you use to measure sizes. All scientists, including astronomers, use the metric system of units because it is well understood worldwide and, more importantly, because it simplifies calculations. If you are not already familiar with the metric system, or if you need a review, study Appendix A before reading on.

The photo in Figure 1-2 is 1 mile across, which equals 1.609 kilometers. You can see that a kilometer (abbreviated km) is a bit under two-thirds of a mile—a short walk across a neighborhood. But when you expand your field of view by a factor of 100, the neighborhood you saw in the previous photo vanishes (■ Figure 1-3). Now your field of view is 160 km wide, and you see cities and towns as patches of gray. Wilmington, Delaware, is visible at the lower right. At this scale, you can see some of the natural features of Earth's surface. The Allegheny Mountains of southern Pennsylvania cross the image in the upper left, and the Susquehanna River flows southeast into Chesapeake Bay. What look like white bumps are a few puffs of clouds.

Figure 1-3 is an infrared photograph in which healthy green leaves and crops show up as red. Human eyes are sensitive to only a narrow range of colors. As you explore the universe in the following chapters, you will learn to use a wide range of other "colors," from X-rays to radio waves, to reveal sights invisible to unaided human eyes.

At the next step in your journey, you can see your entire planet, which is nearly 13,000 km in diameter (■ Figure 1-4). At any particular moment, half of Earth's surface is exposed to sunlight, and half is in darkness. As Earth rotates on its axis, it carries you through sunlight and then through darkness, producing the cycle of day and night. The blurriness you see at the extreme right of the photo is the boundary between day and night—the sunset line. This is a good example of how a photo can give you visual clues to understanding a concept. Special questions called "Learning to Look" at the end of each chapter give you a chance to use your own imagination to connect images with explanations about astronomical objects.

■ **Figure 1-4**

NASA

Enlarge your field of view by a factor of 100, and you see a region 1,600,000 km wide (■ Figure 1-5). Earth is the small blue dot in the center, and the moon, whose diameter is only one-fourth that of Earth, is an even smaller dot along its orbit 380,000 km away.

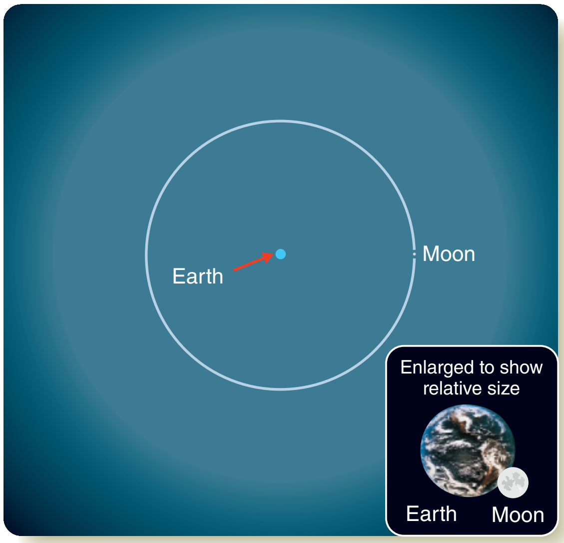
These numbers are so large that it is inconvenient to write them out. Astronomy is sometimes known as the science of big numbers, and soon you will be using numbers much larger than these to discuss the universe. Rather than writing out these numbers as in the previous paragraph, it is more convenient to write them in **scientific notation.** This is nothing more than a simple way to write very big or very small numbers without using lots of zeros. In scientific notation, 380,000 becomes 3.8×10^5. If you are not familiar with scientific notation, read the section on powers of 10 notation in the Appendix. The universe is too big to discuss without using scientific notation.

When you once again enlarge your field of view by a factor of 100, Earth, the moon, and the moon's orbit all lie in the small red box at lower left of ■ Figure 1-6. Now you can see the sun and two other planets that are part of our solar system. Our **solar system** consists of the sun, its family of planets, and some smaller bodies, such as moons and comets.

Earth, Venus, and Mercury are **planets,** small, spherical, nonluminous bodies that orbit a star and shine by reflected light. Venus is about the size of Earth, and Mercury is just over a third of Earth's diameter. On this diagram, they are both too small to be seen as anything but tiny dots. The sun is a **star,** a self-luminous ball of hot gas that generates its own energy. Even though the sun is 109 times larger in diameter than Earth (inset), it too is nothing more than a dot in this diagram.

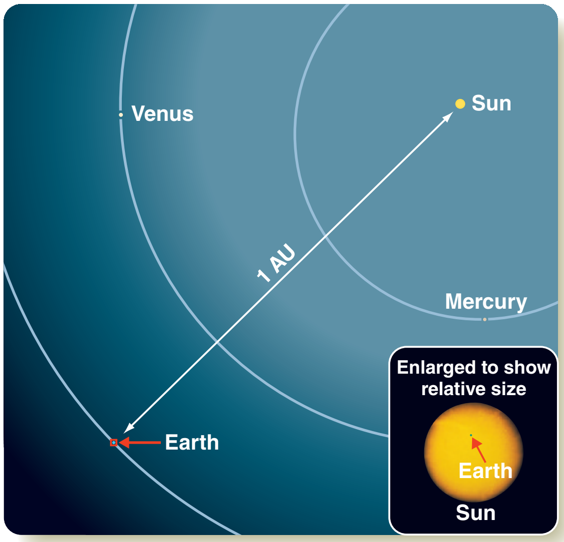
This diagram represents an area with a diameter of 1.6×10^8 km. One way astronomers simplify calculations using large numbers is to define larger units of measurement. For example, the average distance from Earth to the sun is a unit of distance called the **astronomical unit (AU),** which is equal to 1.5×10^8 km. Now you can express the average distance from Venus to the

■ **Figure 1-5**



© Cengage Learning 2014; Images: NASA

■ **Figure 1-6**



© Cengage Learning 2014; Images: NOAO/AURA/NSF

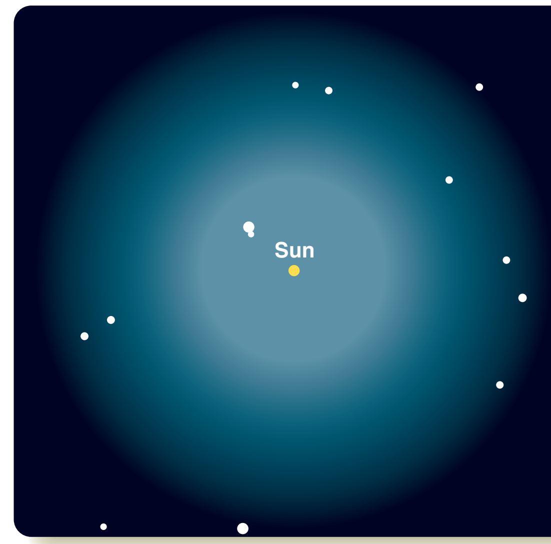
sun as about 0.72 AU and the average distance from Mercury to the sun as about 0.39 AU.

These distances are averages because the orbits of the planets are not perfect circles. This is particularly apparent in the case of Mercury. Its orbit carries it as close to the sun as 0.307 AU and as far away as 0.467 AU. You can see the variation in the distance from Mercury to the sun in Figure 1-6. Earth's orbit is more circular, and its distance from the sun varies by only a few percent.

Enlarge your field of view again, and you can see the entire solar system (■ Figure 1-7). The sun, Mercury, Venus, and Earth lie so close together that you cannot see them separately at this scale, and they are lost in the red square at the center of this diagram. You can see only the brighter, more widely separated objects such as Mars, the next planet outward. Mars lies only 1.5 AU from the sun, but Jupiter, Saturn, Uranus, and Neptune are farther from the sun and so are easier to place in this diagram. They are cold worlds far from the sun's warmth. Light from the sun reaches Earth in only 8 minutes, but it takes over 4 hours to reach Neptune.

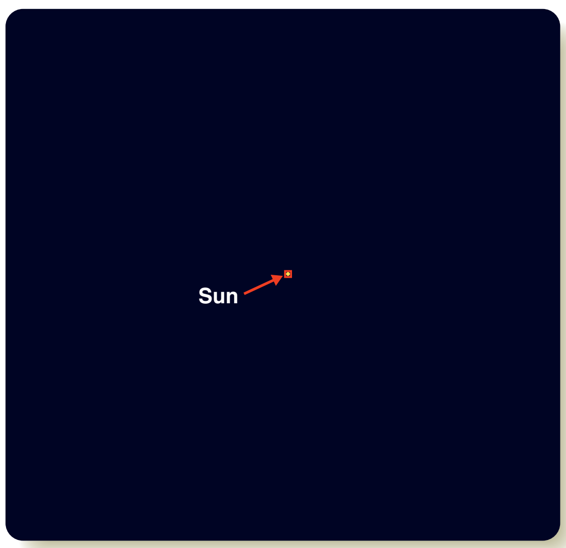
You can remember the order of the plants from the sun outward by remembering a simple sentence: *My Very Educated Mother Just Served Us Noodles.* The first letter of each word reminds you of a planet: Mercury, Venus Earth, Mars Jupiter, Saturn, Uranus, Neptune.

■ **Figure 1-9**



© Cengage Learning 2014

■ **Figure 1-8**



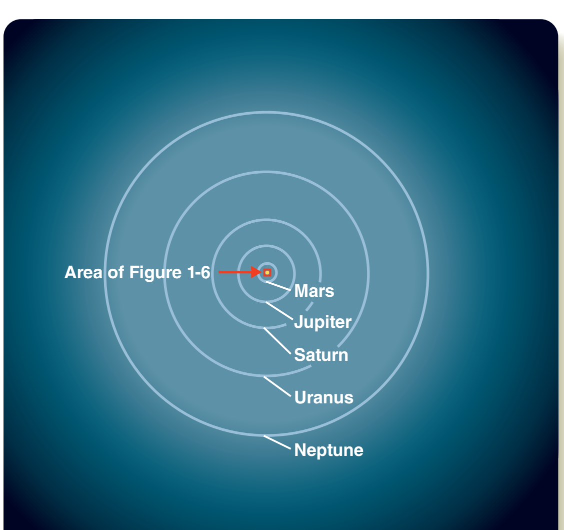
© Cengage Learning 2014

When you again enlarge your field of view by a factor of 100, the solar system vanishes (■ Figure 1-8). The sun is only a point of light, and all the planets and their orbits are now crowded into the small red square at the center. The planets are too small and too faint to be visible so near the brilliance of the sun.

Nor are any stars visible except for the sun. The sun is a fairly typical star, and it seems to be located in a fairly average neighborhood in the universe. Although there are many billions of stars like the sun, none are close enough to be visible in this diagram, which shows a region only 11,000 AU in diameter. Stars are typically separated by distances about 10 times larger than that.

In ■ Figure 1-9, your field of view has expanded to a diameter of a bit over 1 million AU. The sun is at the center, and at this scale you can see a few of the nearest stars. These stars are so distant that it is not convenient to give their distances in astronomical units. To express distances so large, astronomers define a new unit of distance, the light-year. One **light-year (ly)** is the distance that light travels in one year, roughly 10^{13} km or 63,000 AU. It is a **Common Misconception** that a light-year is a unit of time, and you can sometimes hear the term misused in science fiction movies and TV shows. The next time you hear someone say, "It will take me light-years to finish my history paper," you can tell

■ **Figure 1-7**



© Cengage Learning 2014

that person that a light-year is a distance, not a time. The diameter of your field of view in Figure 1-9 is 17 ly.

Another **Common Misconception** is that stars look like disks when seen through a telescope. Although stars are roughly the same size as the sun, they are so far away that astronomers cannot see them as anything but points of light. Even the closest star to the sun—Proxima Centauri, only 4.2 ly from Earth—looks like a point of light through even the biggest telescopes on Earth. Furthermore, planets that circle other stars are much too small, too faint, and too close to the glare of their star to be easily visible. Astronomers have used indirect methods to detect over 800 planets orbiting other stars, but only a few have been photographed directly, and even those show up as nothing more than faint points of light.

Figure 1-9 follows the astronomical custom of making the sizes of the dots represent not the sizes of the stars but their brightnesses. This is how star images are recorded on photographs. Bright stars make larger spots on a photograph than faint stars, so the size of a star image in a photograph tells you not how big the star is but only how bright it looks.

In ■ Figure 1-10, you expand your field of view by another factor of 100, and the sun and its neighboring stars vanish into the background of thousands of other stars. The field of view is now 1700 ly in diameter. Of course, no one has ever journeyed thousands of light-years from Earth to look back and photograph the solar neighborhood, so this is a representative photograph of the sky. The sun is a relatively faint star that would not be easily located in a photo at this scale.

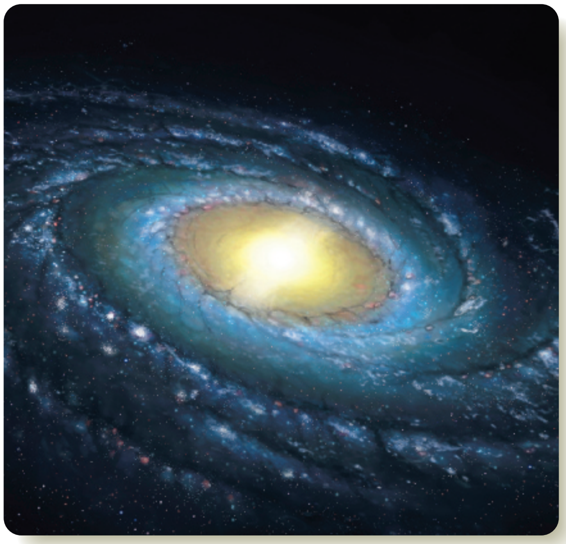
If you again expand your field of view by a factor of 100, you see our galaxy, a disk of stars about 80,000 ly in diameter (■ Figure 1-11). A **galaxy** is a great cloud of stars, gas, and dust held together by the combined gravity of all of its matter. Galaxies range from 1500 to over 300,000 ly in diameter, and some contain over 100 billion stars. In the night sky, you can see our galaxy as a great, cloudy wheel of stars surrounding us and ringing the sky. This band of stars is known as the **Milky Way,** and our galaxy is called the **Milky Way Galaxy.**

How does anyone know what our galaxy looks like if no one can leave it and look back? Astronomers use evidence to guide their theories as they imagine what the Milky Way looks like. Artists can then use those scientific descriptions to create a painting. Many images in this book are artists' renderings of objects and events that are too big or too dim to see clearly, emit energy your eyes cannot detect, or hap-pen too slowly or too rapidly for humans to sense. These images are not just guesses; they are scientifically based illustrations guided by the best information astronomers can gather.

■ **Figure 1-10**

NOAO/AURA/NSF

■ **Figure 1-11**



© Mark Garlick/Space-art.com

■ **Figure 1-12**

© Cengage Learning 2014

As you explore, notice how astronomers use science to imagine, understand, and depict cosmic events.

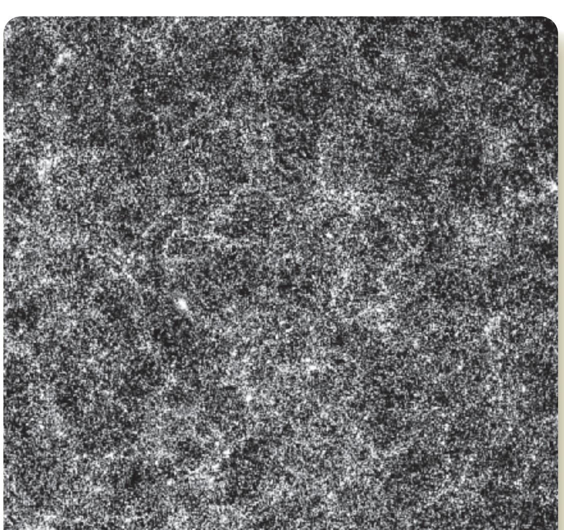
The artist's conception of the Milky Way reproduced in Figure 1-11 shows that our galaxy, like many others, has graceful **spiral arms** winding outward through its disk. In a later chapter, you will learn that the spiral arms are places where stars are formed from clouds of gas and dust. Our own sun was born in one of these spiral arms; if you could see it in this picture, it would be in the disk of the galaxy about two-thirds of the way out from the center.

Ours is a fairly large galaxy. Only a century ago astronomers thought it was the entire universe—an island cloud of stars in an otherwise empty vastness. Now they know that our galaxy is not unique; it is only one of many billions of galaxies scattered throughout the universe.

You can see a few of these other galaxies when you expand your field of view by another factor of 100 (■ Figure 1-12). Our galaxy appears as a tiny luminous speck surrounded by other specks in a region 17 million light-years in diameter. Each speck represents a galaxy. Notice that our galaxy is part of a cluster of a few dozen galaxies. Galaxies are commonly grouped together in such clusters. Some galaxies have beautiful spiral patterns like our own galaxy, but others do not. Some are strangely distorted. In a later chapter, you will learn what produces these differences among the galaxies.

Now is a chance for you to correct another **Common Misconception.** People often say "galaxy" when they mean "solar system," and they sometimes confuse both terms with "universe." Your cosmic zoom has shown you the difference. The **solar system** is the sun and its planets, including Earth. Our **galaxy** contains our solar system plus billions of other stars and whatever planets orbit around them. The **universe** includes everything: billions of galaxies, each containing billions of stars and, presumably, billions of planetary systems.

If you expand your field of view one more time, you can see that clusters of galaxies are connected in a vast network (■ Figure 1-13). Clusters are grouped into superclusters—clusters of clusters—and the superclusters are linked to form long filaments and walls outlining nearly empty voids. These filaments and walls appear to be the largest structures in the universe. Were you to expand your field of view another time, you would probably see a uniform fog of filaments and walls. When you puzzle over the origin of these structures, you are at the frontier of human knowledge.



Based on data from M. Seldner, B. L. Siebers, E. J. Groth, and P. J. E. Peebles, *Astronomical Journal 82* [1977].

■ **Figure 1-13**

1-2 When Is Now?

Now that you have an idea where you are in space, you need to know where you are in time. The stars have shone for billions of years before the first human looked up and wondered what they were. To get a sense of your place in time, all you need is a long red ribbon.

Imagine stretching that ribbon down the center of an American football field from goal line to goal line, a distance of 100 yards (about 91 meters), as shown on the inside front cover of this book. Further, imagine that one end of the ribbon represents *today* and the other end represents the beginning of the universe—the moment of beginning that astronomers call the *big bang.* In a later chapter, "Modern Cosmology," you will learn about the big bang, and you will see evidence that the universe is approximately 14 billion years old. Your long red ribbon represents 14 billion years, the entire history of the universe.

Imagine beginning at the goal line labeled *Big Bang* and replaying the entire history of the universe as you walk along your ribbon toward the goal line labeled *today.* Observations tell astronomers that the big bang filled the entire universe with hot, glowing gas, but as the gas cooled and dimmed the universe went dark. All that happened along the first half inch of the ribbon. There was no light for the next 400 million years, until gravity was able to pull some of the gas together to form the first stars. That seems like a lot of years, but if you stick a little flag beside the ribbon to mark the birth of the first stars, it would be not quite 3 yards from the goal line where the universe began.

You must walk only about 5 yards along the ribbon before galaxies formed in large numbers. Our home galaxy would be one of those taking shape. By the time you cross the 50-yard line, the universe is full of galaxies, but the sun and Earth have not formed yet. You must walk past the 50-yard line down to the 35-yard line before you can finally stick a flag beside the ribbon to mark the formation of the sun and planets—our solar system—4.6 billion years ago, about 9 billion years after the big bang.

You must carry your flags a few yards further to the 29-yard line to mark the appearance of the first life on Earth—microscopic creatures in the oceans—and you have to walk all the way to the 3-yard line before you can mark the emergence of life on land. Your dinosaur flag goes just inside the 2-yard line. Dinosaurs go extinct as you pass the one-half-yard line.

What about people? The first humanlike creatures appeared on Earth about 4 million years ago, so you can put a little flag for the first humans only about an inch from the goal line labeled *today.* Civilization, the building of cities, began about 10,000 years ago, so you have to try to fit that flag in only 0.0026 inch from the goal line. That's half the thickness of a sheet of paper. Compare the history of human civilization with the history of the universe. Every war you have ever heard of, every person whose name is recorded, every structure ever built from Stonehenge to the building you are in right now fits into that 0.0026 inch.

Humanity is very new to the universe. Our civilization on Earth has existed for only a flicker of an eyeblink in the history of the universe. As you will discover in the chapters that follow, only in the last hundred years or so have astronomers begun to understand where we are in space and in time.

1-3 Why Study Astronomy?

Your exploration of the universe will help you answer two fundamental questions:

What are we?
How do we know?

The question "What are we?" is the first organizing theme of this book. Astronomy is important to you because it will tell you what you are. Notice that the question is not "*Who* are we?" If you want to know who we are, you may want to talk to a sociologist, theologian, paleontologist, artist, or poet. "*What* are we?" is a fundamentally different question.

As you study astronomy, you will learn how you fit into the history of the universe. You will learn that the atoms in your body had their first birthday in the big bang when the universe began. Those atoms have been cooked and remade inside generations of stars, and now, after billions of years, they are inside you. Where will they be in another billion years? This is a story everyone should know, and astronomy is the only course on campus that can tell you that story.

Every chapter in this book ends with a short segment titled "What Are We?" This summary shows how the astronomy in the chapter relates to your role in the story of the universe.

The question "How do we know?" is the second organizing theme of this book. It is a question you should ask yourself whenever you encounter statements made by so-called experts in any field. Should you swallow a diet supplement recommended by a TV star? Should you vote for a candidate who warns of a climate crisis? To understand the world around you and to make wise decisions for yourself, for your family, and for your nation, you need to understand how science works.

You can use astronomy as a case study in science. In every chapter of this book, you will find short essays titled "How Do We Know?" They are designed to help you think not about *what* is known but about *how* it is known. To do that, they will explain different aspects of scientific reasoning and in that way help you understand how scientists know about the natural world.

Over the last four centuries, scientists have developed a way to understand nature by comparing hypotheses with evidence, a process that has been called the **scientific method** (**How Do We Know? 1-1**). As you read about exploding stars, colliding galaxies, and alien planets in the following chapters, you will see astronomers using the scientific method over and over. The universe is very big, but it is described by a small set of rules, and we humans have found a way to figure out the rules—a method called *science.*

How Do We Know? 1-1

The So-Called Scientific Method

How do scientists learn about nature? You have probably heard of the **scientific method** as the process by which scientists form hypotheses and test them against evidence gathered by experiment or observation. Scientists use the scientific method all the time, and it is critically important, but they rarely think of it at all, and they certainly don't think of it as a numbered list of steps. It is such an ingrained way of thinking and understanding nature that it is almost invisible to the people who use it most.

Scientists try to form hypotheses that explain how nature works. If a hypothesis is contradicted by evidence from experiments or observations, it must be revised or discarded. If a hypothesis is confirmed, it must be tested further. In that very general way, the scientific method is a way of testing and refining ideas to better describe how nature works.

For example, Gregor Mendel (1822–1884) was an Austrian abbot who liked plants. He formed a hypothesis that offspring usually inherit traits from their parents not as a smooth blend, as most scientists of the time believed, but according to strict mathematical rules. Mendel cultivated and tested over 28,000 pea plants, noting which produced smooth peas and which produced wrinkled peas and how that trait was inherited by successive generations. His study of pea plants confirmed his hypothesis and allowed the development of a series of laws of inheritance. Although the importance of his work was not recognized in his lifetime, Mendel is now called the "father of modern genetics."

The scientific method is not a simple, mechanical way of grinding facts into understanding. It is, in fact, a combination of many ways of analyzing information, finding relationships, and creating new ideas. A scientist needs insight and ingenuity to form and test a good hypothesis. Scientists use the scientific method almost automatically, forming, testing, revising, and discarding hypotheses almost minute by minute as they discuss a new idea. Sometimes, however, a scientist will spend years studying a single promising hypothesis. The so-called scientific method is a way of thinking and a way of knowing about nature. The "How Do We Know?" essays in the chapters that follow will introduce you to some of those methods.

Inspirestock/Jupiterimages

Whether peas are wrinkled or smooth is an inherited trait.

What Are We? Part of the Story

Astronomy will give you perspective on what it means to be here on Earth. This chapter has helped you locate yourself in space and time. Once you realize how vast our universe is, Earth seems quite small. People on the other side of the world seem like neighbors. And, in the entire history of the universe, the human story is only the blink of an eye. This may seem humbling at first, but you can be proud of how much we humans have understood in such a short time.

Not only does astronomy locate you in space and time, it places you in the physical processes that govern the universe. Gravity and atoms work together to make stars, light the universe, generate energy, and create the chemical elements in your body. The chapters that follow will show how you fit into that cosmic process.

Although you are very small and your kind have existed in the universe for only a short time, you are an important part of something very large and very beautiful.

Study and Review

Summary

- You surveyed the universe by taking a cosmic zoom in which each **field of view (p. 2)** was 100 times wider than the previous field of view.
- Astronomers use the metric system because it simplifies calculations and use **scientific notation (p. 3)** for very large or very small numbers.
- You live on a **planet (p. 3),** Earth, which orbits our **star (p. 3),** the sun, once a year. As Earth rotates once a day, you see the sun rise and set.
- The moon is only one-fourth the diameter of Earth, but the sun is 109 times larger in diameter than Earth—a typical size for a star.
- The **solar system (p. 3)** includes the sun at the center, all of the planets that orbit around it—Mercury, Venus, Earth, Mars, Jupiter, Saturn, Uranus, and Neptune—plus the moons of the planets and all other objects bound to the sun by its gravity.
- The **astronomical unit (AU) (p. 3)** is the average distance from Earth to the sun. Mars, for example, orbits 1.5 AU from the sun. The **light-year (ly) (p. 4)** is the distance light can travel in one year. The nearest star is 4.2 ly from the sun.
- Many stars seem to have planets, but such small, distant worlds are difficult to detect. Nevertheless, over 400 have been found after almost 20 years of searching by astronomers. So far only a few of those planets have been determined to be Earth-like in terms of size and temperature.
- The **Milky Way (p. 5),** the hazy band of light that encircles the sky, is the **Milky Way Galaxy (p. 5)** seen from inside. The sun is just one out of the billions of stars that fill the Milky Way Galaxy.
- **Galaxies (p. 5)** contain many billions of stars. Our galaxy is about 80,000 ly in diameter and contains over 100 billion stars.
- Some galaxies, including our own, have graceful **spiral arms (p. 6)** bright with stars, but some galaxies are plain clouds of stars.
- The **solar system (p. 6)** consists of the sun plus eight planets, including Earth. Our **galaxy (p. 6)** contains our solar system plus billions of other stars and whatever planets orbit around them. The **universe (p. 6)** includes everything that there is: billions of galaxies, each containing billions of stars and, presumably, billions of planetary systems.
- Our galaxy is just one of billions of galaxies that fill the universe in great clusters, clouds, filaments, and walls—the largest structures in the universe.
- The universe began about 14 billion years ago in an event called the big bang, which filled the universe with hot gas.
- The hot gas cooled, the first galaxies began to form, and stars began to shine only about 400 million years after the big bang.
- The sun and planets of our solar system formed about 4.6 billion years ago.
- Life began in Earth's oceans soon after Earth formed but did not emerge onto land until only 400 million years ago. Dinosaurs evolved not long ago and went extinct only 65 million years ago.
- Humanlike creatures developed on Earth only about 4 million years ago, and human civilizations developed only about 10,000 years ago.
- Although astronomy seems to be about stars and planets, it describes the universe in which you live, so it is really about you. Astronomy helps you answer the question, "What are we?"
- As you study astronomy, you should ask "How do we know?" and that will help you understand how science gives us a way to understand nature.
- In its simplest outline, science follows the **scientific method (p. 7),** by which scientists test hypotheses against evidence from experiments and observations. This method is a powerful way to learn about nature.

Review Questions

1. What is the largest dimension of which you have personal knowledge? Have you run a mile? Hiked 10 miles? Run a marathon?
2. What is the difference between our solar system, our galaxy, and the universe?
3. Why are light-years more convenient than miles, kilometers, or astronomical units for measuring certain distances?
4. Why is it difficult to detect planets orbiting other stars?
5. What does the size of the star image in a photograph tell you?
6. What is the difference between the Milky Way and the Milky Way Galaxy?
7. What are the largest known structures in the universe?
8. How does astronomy help answer the question "What are we?"
9. **How Do We Know?** How does the scientific method give scientists a way to know about nature?

Discussion Questions

1. Do you think you have a duty to know the astronomy described in this chapter? Can you think of ways this knowledge helps you enjoy a richer life and be a better citizen?
2. How is a statement in a political campaign speech different from a statement in a scientific discussion? Find examples in newspapers, magazines, and this book.

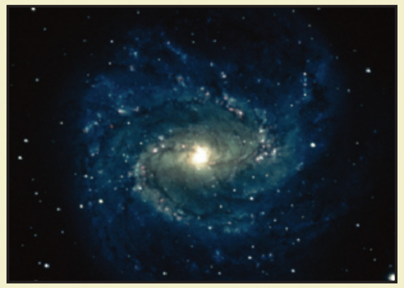
Problems

1. The diameter of Earth across the equator is 7928 miles. If a mile equals 1.609 km, what is Earth's diameter in kilometers? In centimeters?
2. The diameter of the moon across its equator is 3476 kilometers. If a kilometer equals 0.6214 miles, what is the moon's diameter in miles?
3. One astronomical unit is about 1.50×10^8 km. Explain why this is the same as 150×10^6 km.
4. Venus orbits 0.72 AU from the sun. What is that distance in kilometers?
5. Light from the sun takes 8 minutes to reach Earth. How long does it take to reach Mars?
6. The sun is almost 400 times farther from Earth than is the moon. How long does light from the moon take to reach Earth?
7. If the speed of light is 3.00×10^5 km/s, how many kilometers is 1 light-year? How many meters? (*Note:* One year contains 3.16×10^7 s.)
8. How long does it take light to cross the diameter of our Milky Way Galaxy?
9. The nearest large galaxy to our own is about 2.5 million light-years away. How many meters is that?
10. How many galaxies like our own would it take laid edge-to-edge to reach the nearest galaxy? (*Hint:* See Problem 9.)

Learning to Look

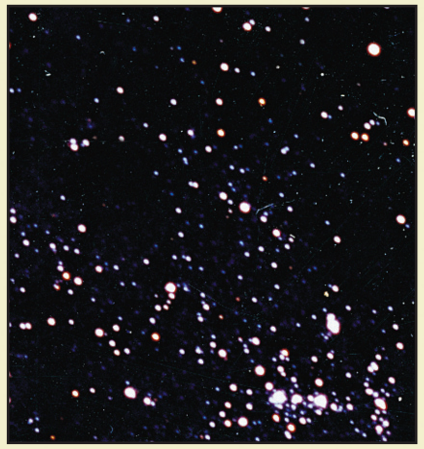
1. In Figure 1-4, the division between daylight and darkness is at the right on the globe of Earth. How do you know this is the sunset line and not the sunrise line?
2. Look at Figure 1-6. How can you tell that Mercury does not follow a circular orbit?
3. Of the objects listed here, which would be contained inside the object shown in the photograph at the right? Which would contain the object in the photo?

   stars
   planets
   galaxy clusters
   filaments
   spiral arms



Bill Schoening/NOAO/AURA/NSF

4. In the photograph shown here, which stars are brightest, and which are faintest? How can you tell? Why can't you tell which stars in this photograph are biggest or which have planets?



NOAO/AURA/NSF

Great Debates

1. *Moon Ownership.* The Outer Space Treaty was signed by the United States, the United Kingdom, and the Soviet Union on October 10, 1967. The treaty expressly forbids any government from claiming ownership of a celestial resource such as the moon. On July 21, 1969, the United States was the first to land and plant its flag on the moon. Did the United States violate the treaty by "planting the flag," an act that has historically claimed ownership? If so, does the United States own the moon? The United Nations Treaties and Principles on Space Law can be found at www.oosa.unvienna.org/oosa/SpaceLaw/treaties.html.
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

2. *Another 2012?* In December 2012, the world was supposed to come to an end according to various beliefs. Suppose a new prophecy is found that predicts a new ending date to the world, and the new date is two years from today. Your neighbor decides to build an underground bunker in his backyard. The bunker can house only sixty people, and only those that fund the bunker will be allowed to enter. Your neighbor creates an online community blog, tweet, and posts to advertise. What do you do? Do you voice your concerns on the blog, tweets, and posts, knowing your name can be seen? What do you say? Do you vote yes to join and fund the bunker?
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

3. *Anyone Out There?* Should we humans send out signals to anyone who is listening and advertise our location in space? Supposing that listeners exist, will they be able to find us based upon the signal they receive from us, and should we humans worry?
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

4. *What Are We?* Based on the age of the universe, the age of the solar system, and the age of humanlike creatures, would you consider humans as the most intelligent beings in the universe? How would you rate the level of human intelligence compared to other intelligence that may exist in the universe? Very low? Low? Intermediate? High? Very high? The highest?
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

5. *How Do You Know?* Should you believe everything written in this textbook? If information from a television show or the Internet disagrees with what you read in this book, which would you believe? Is "believe" a good word to use for a scientific understanding? If you disagree with the information from the television, Internet, and text, and instead you draw a conclusion based on your own knowledge, how do you know you are correct? Could you be biased in your conclusions?
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

CengageNOW

CengageNOW™ is an integrated, online learning system that gives you 24/7 access to Study Tools and assignments. Working at your own pace, or within a schedule set up by your instructor, you can read the textbook, take quizzes, and track your grades in an easy-to-use, personalized online environment that you manage to best suit your needs. You can also master key concepts and prepare for exams with CengageNOW's Personalized Study—a diagnostic program plus study plan—preloaded with an integrated eBook and other text-specific material.

CengageNOW Virtual Astronomy Laboratories 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Laboratories 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Laboratories. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Laboratory 1.1: Measurement and Unit Conversion

A vital characteristic of scientific knowledge is that it is supported by evidence. A scientific statement is more than an opinion or a speculation because it has been tested objectively against reality. When you think of evidence, you probably think of criminal investigations in which detectives collect fingerprints and eyewitness accounts. In science, evidence most often consists of measurements.

One of the most difficult but crucial measurements in astronomy is the measurement of distance. If you don't know the distance to an object, you know only how large and how bright it seems from your standpoint on Earth. You don't know its real size, its total energy output, or any other absolute quantities.

As you learned in this chapter, the distances to astronomical objects are . . . well, astronomical. Use of scientific notation (also called "powers of 10 notation") simplifies writing very large and very small numbers. For example, the average distance between Earth and the sun is 150,000,000 km. Writing this number as 1.5×10^{8} km is much easier. The wavelength of light that your eye sees as yellow is about 0.00000058 m. In scientific notation that becomes 5.8×10^{-7} m.

As you also learned in this chapter, inventing new units of measure such as the AU and the parsec is another way to make astronomical distances easier to deal with. Conversion factors between various common units can be found in this book's Appendix Tables A-3, A-4, and A-6.

Sections 1 and 2 of Virtual Astronomy Laboratory 1, "Measurement and Unit Conversion," help you practice techniques of recording measurements, converting units, and using scientific notation. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

2

A User's Guide to the Sky

Guidepost

The previous chapter took you on a cosmic zoom through space and time. That quick preview set the stage for the drama to come. In this chapter, you restart your explorations by viewing the sky from Earth with your own unaided eyes. As you do, consider three important questions:

- **How are names chosen for constellations and stars?**
- **How can you compare the brightness of the stars?**
- **How does the sky appear to move as Earth rotates?**

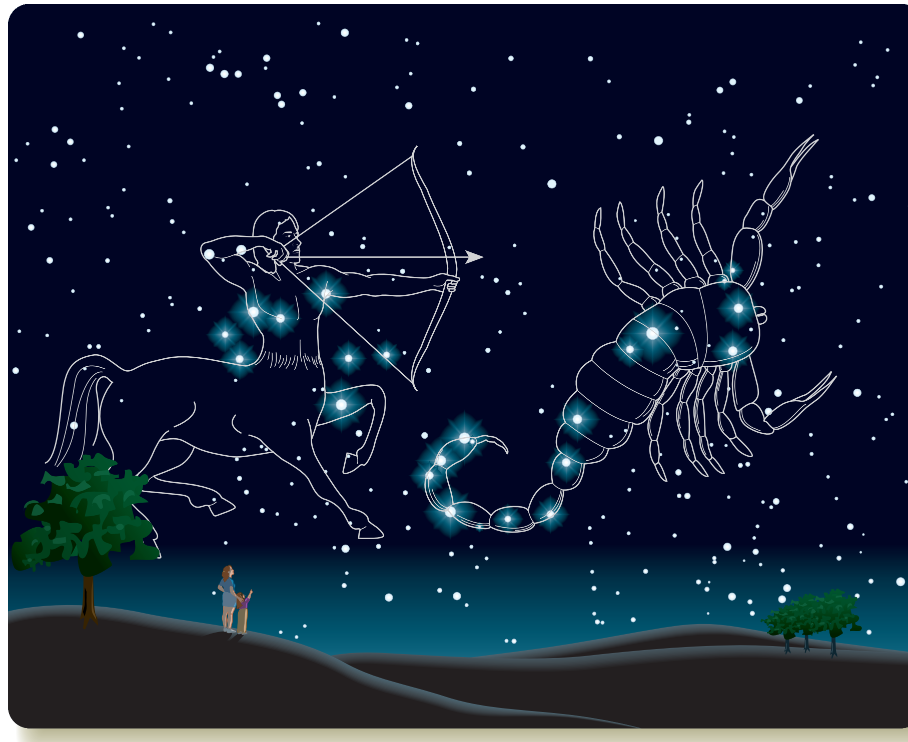
As you study the sky and its apparent daily motions, you will learn to imagine Earth as a planet rotating on its axis. The next chapter will introduce you to celestial cycles caused by the monthly revolution of the moon around Earth and the yearly revolution of Earth around the sun.

A nearly 3-hour exposure of Southern Hemisphere circumpolar star trails. There is no bright star marking the south celestial pole, but star trails for the Southern Cross can be seen at center right.

Laurent Fournet, Jean-Philippe Olive, Club Astro Astrium, Toulouse,France

The Southern Cross I saw every night abeam. The sun every morning came up astern; every evening it went down ahead. I wished for no other compass to guide me, for these were true.

CAPTAIN JOSHUA SLOCUM
SAILING ALONE AROUND THE WORLD



© Cengage Learning 2014

Figure 2-1

The constellations are an ancient heritage handed down for thousands of years as celebrations of great heroes and mythical creatures. Here Sagittarius and Scorpius hang above the southern horizon.

THE NIGHT SKY is the rest of the universe as seen from our planet. When you look up at the stars, you are looking out through a layer of air only a little more than a hundred kilometers deep. Beyond that, space is nearly empty, and the stars are scattered light-years apart.

As you read this chapter, keep in mind that you live on a planet in the midst of these scattered stars. Because Earth turns on its axis once a day, the sky appears to revolve around you in a daily cycle. Not only does the sun rise in the eastern part of the sky and set in the western part, but so do the stars.

2-1 The Stars

ON A DARK NIGHT far from city lights, you can see a few thousand stars. Long ago, humans organized what they saw by naming stars and groups of stars. Some of those names survive today.

Constellations

All around the world, ancient cultures celebrated heroes, gods, and mythical beasts by giving their names to groups of stars—**constellations** (■ Figure 2-1). You should not be surprised that the star patterns do not look like the creatures they represent any more than Columbus, Ohio, looks like Christopher Columbus. The constellations simply celebrate the most important mythical figures in each culture. The oldest constellations named by Western cultures originated in Assyria over 3000 years ago, and others were added by Babylonian and Greek astronomers during the classical age. Of these ancient constellations, 48 are still in use.

To the ancients, a constellation was a loose grouping of stars. Many of the fainter stars were not included in any constellation, and the stars of the southern sky not visible to the ancient astronomers of northern latitudes were not grouped into constellations. Constellation boundaries, when they were defined at all, were only approximate (■ Figure 2-2a), so a star like Alpheratz could be thought of as part of Pegasus or part of Andromeda. To correct these gaps and ambiguities, astronomers have added 40 modern constellations, and in 1928 the International Astronomical Union established 88 official constellations with clearly defined boundaries (Figure 2-2b). Consequently, a constellation now represents not a group of stars but an area of the sky, and any star within the region belongs to one and only one constellation. Alpheratz belongs to Andromeda.

In addition to the 88 official constellations, the sky contains a number of less formally defined groupings called **asterisms**. The Big Dipper, for example, is a well-known asterism that is part of the constellation Ursa Major (the Great Bear). Another asterism is the Great Square of Pegasus (Figure 2-2b), which includes three stars from Pegasus plus Alpheratz from Andromeda. The star charts at the end of this book will introduce you to the brighter constellations and asterisms.

Although constellations and asterisms are groups of stars that appear close together in the sky, it is important to remember that most are made up of stars that are not physically associated with one another. Some stars may be many times farther away than others and moving through space in different directions.

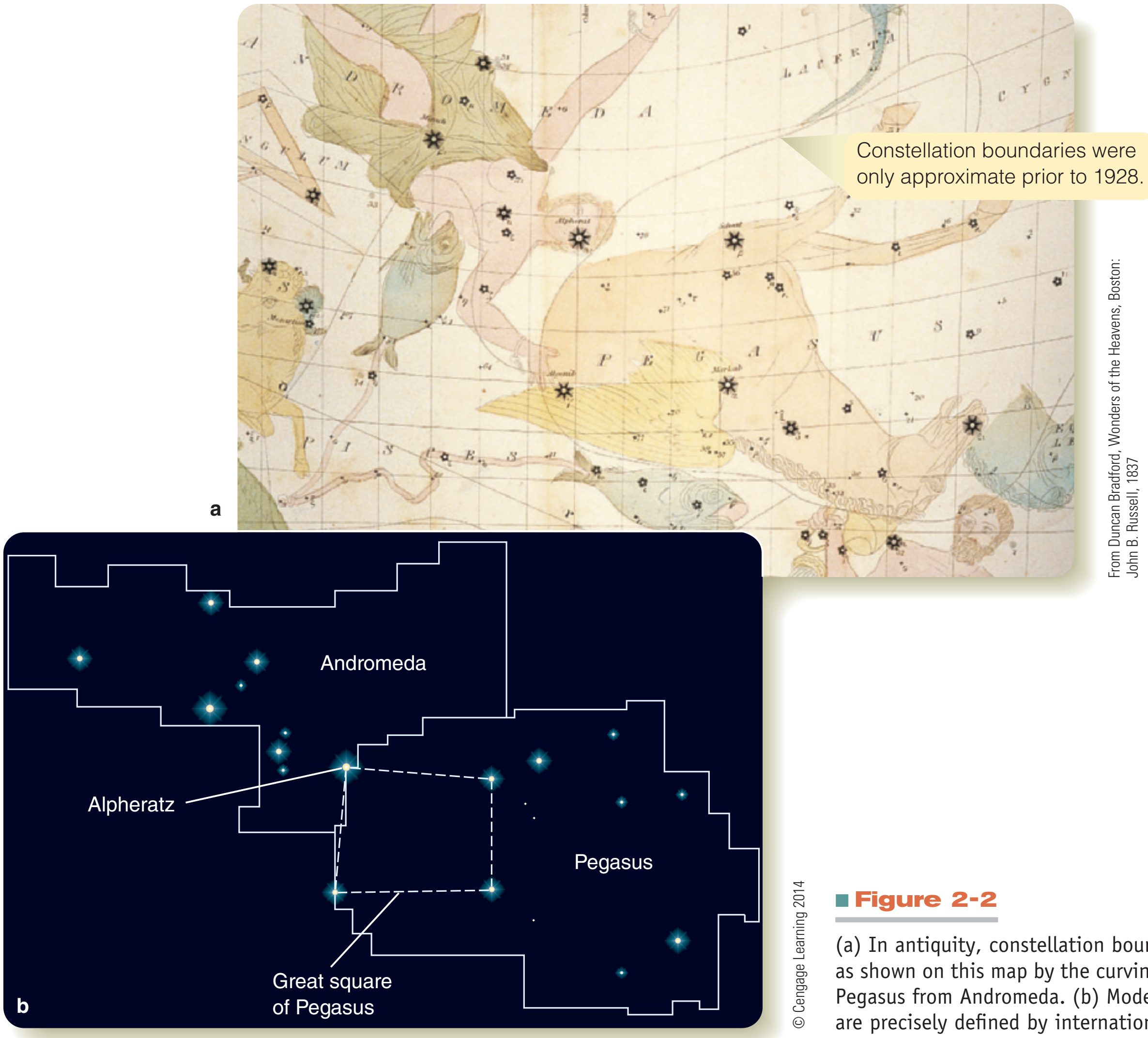


Figure 2-2

(a) In antiquity, constellation boundaries were poorly defined, as shown on this map by the curving dotted lines that separate Pegasus from Andromeda. (b) Modern constellation boundaries are precisely defined by international agreement.

The only thing they have in common is that they happen to lie in approximately the same direction from Earth (Figure 2-3).

The Names of the Stars

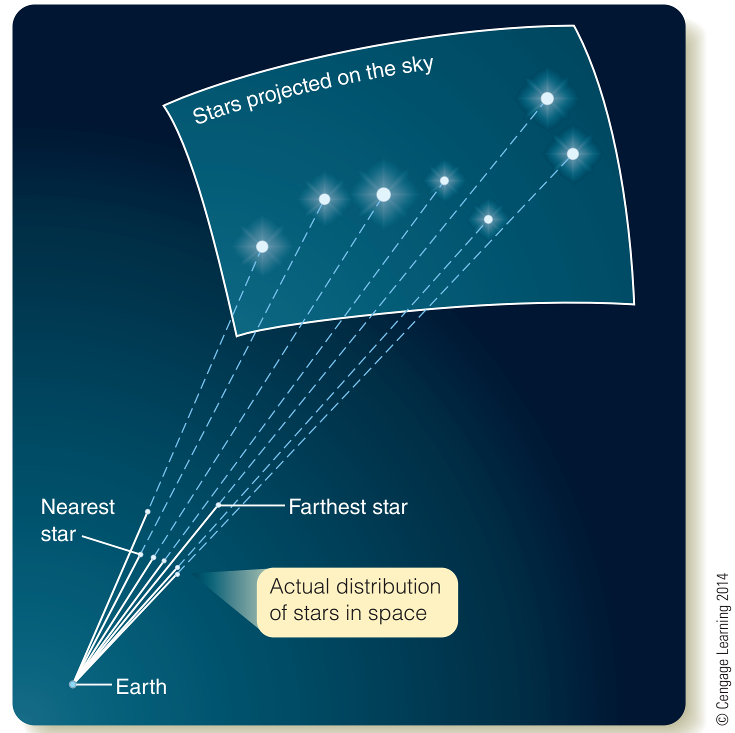
In addition to naming groups of stars, ancient astronomers gave names to the brightest individual stars. Modern astronomers still use many of those ancient names. Although the constellation names came from Greek translated into Latin—the language of science until the 19th century—most star names come from ancient Arabic, though much altered by the passing centuries. The name of Betelgeuse, the bright orange star in Orion, for example, comes from the Arabic *yad al jawza,* meaning "shoulder of Jawza (Orion)." Names such as Sirius ("Scorcher") and Aldebaran ("The Follower [of the Pleiades]") are beautiful additions to the mythology of the sky.

Naming individual stars is not very helpful because you can see thousands of them. How many names could you remember? Also, a simple name gives you no clues to the location of the star in the sky or to its brightness. A more useful way to identify stars is to assign letters to the bright stars in a constellation in approximate order of brightness. Astronomers use the Greek alphabet for this purpose. Thus, the brightest star in a constellation is usually designated alpha, the second brightest beta, and so on. Often the name of the Greek letter is spelled out, as in "alpha," but sometimes the actual Greek letter is used, especially in charts. You will find the Greek alphabet in Appendix A. For many constellations, the letters follow the order of brightness, but some constellations, by tradition, mistake, or the personal preferences of early chart makers, are exceptions (Figure 2-4).

To identify a star by its Greek-letter designation, you would give the Greek letter followed by the possessive (genitive) form of the constellation name; for example, the brightest star in the constellation Canis Major is alpha Canis Majoris, which can also be written α Canis Majoris. This both identifies the star and the constellation and gives a clue to the relative brightness of the star. Compare this with the ancient name for this star, Sirius, which tells you nothing about location or brightness.

■ **Figure 2-3**

You see the Big Dipper in the sky because you are looking through a group of stars scattered through space at different distances from Earth. You see them as if they were projected on a screen, and they form the shape of the Dipper.



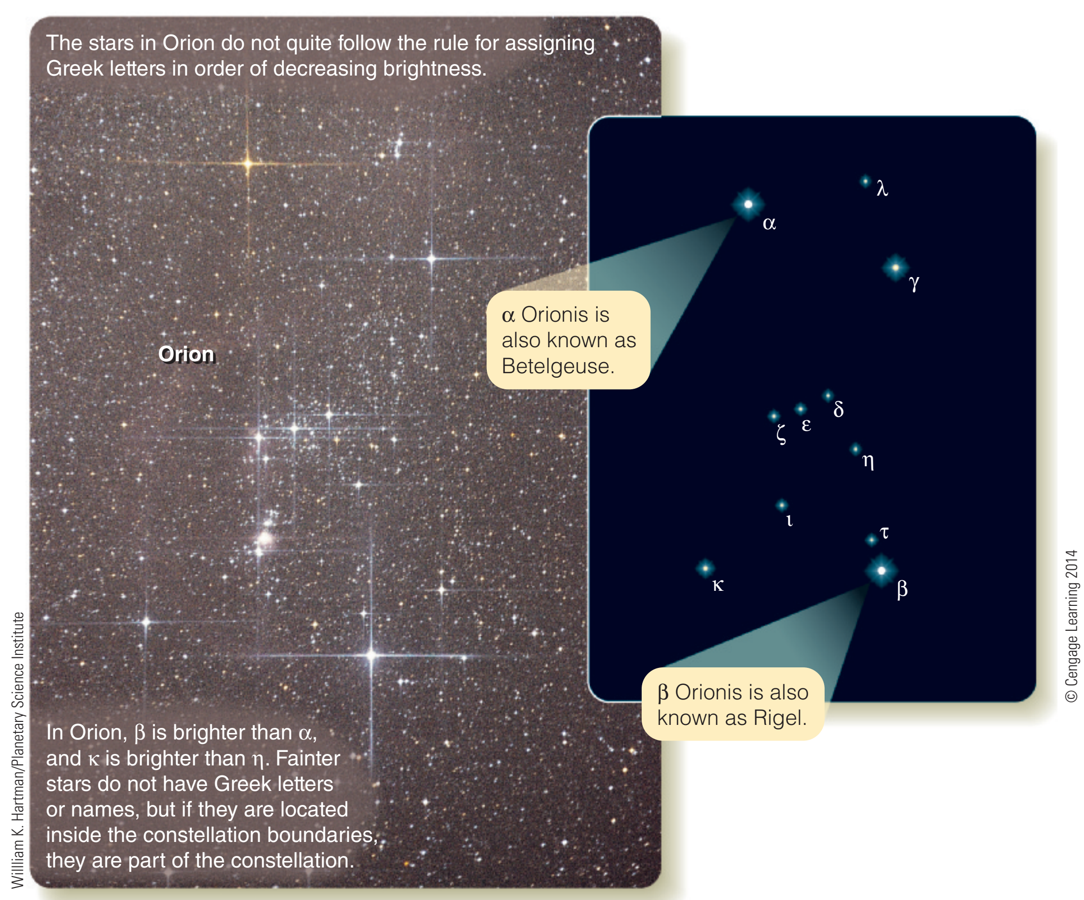
Favorite Stars

It is fun to know the names of the brighter stars, but they are more than points of light in the sky. They are glowing spheres of gas much like the sun, each with its unique characteristics. ■ Figure 2-5 identifies eight bright stars that you can adopt as Favorite Stars. As you study astronomy you will discover their peculiar personalities and enjoy finding them in the evening sky. You will learn, for example, that Betelgeuse is not just an orange point of light but is an aging, cool star over 800 times larger than the sun. As you learn more in later chapters, you may want to add more Favorite Stars to your list.

You can use the star charts at the end of this book to help you locate these Favorite Stars. You can see Polaris year round, but Sirius, Betelgeuse, Rigel, and Aldebaran are in the winter sky. Spica is a summer star, and Vega is visible evenings in later summer and fall. Alpha Centauri, only 4 ly away, is the nearest star to the sun, and you will have to travel as far south as southern Florida to glimpse it above the southern horizon.

The Brightness of Stars

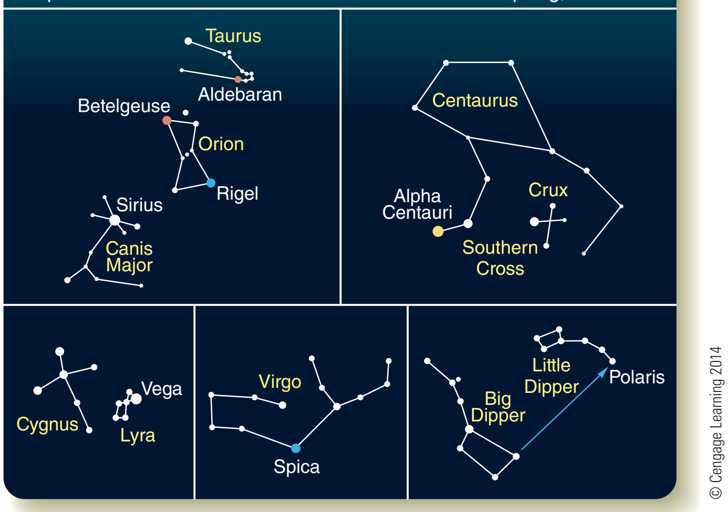
Besides naming individual stars, astronomers need a way to describe their brightness. Astronomers measure the brightness of stars using the **magnitude scale**, a system that first appeared in the writings of the ancient astronomer Claudius Ptolemy about the year 140. The system probably originated even earlier, and most astronomers attribute it to the Greek astronomer Hipparchus (about 190–120 BCE). Hipparchus compiled the first known star catalog, and he may have used the magnitude system in that catalog.



■ **Figure 2-4**

Stars in a constellation can be identified by Greek letters and also by names derived from Arabic. The spikes on the star images in the photograph were produced by the optics in the camera.

Sirius	Brightest star in the sky	Winter
Betelgeuse	Bright red star in Orion	Winter
Rigel	Bright blue star in Orion	Winter
Aldebaran	Red eye of Taurus the Bull	Winter
Polaris	The North Star	Year round
Vega	Bright star overhead	Summer
Spica	Bright southern star	Summer
Alpha Centauri	Nearest star to the sun	Spring, far south



■ **Figure 2-5**

Favorite Stars: Locate these bright stars in the sky and learn about their interesting characteristics.

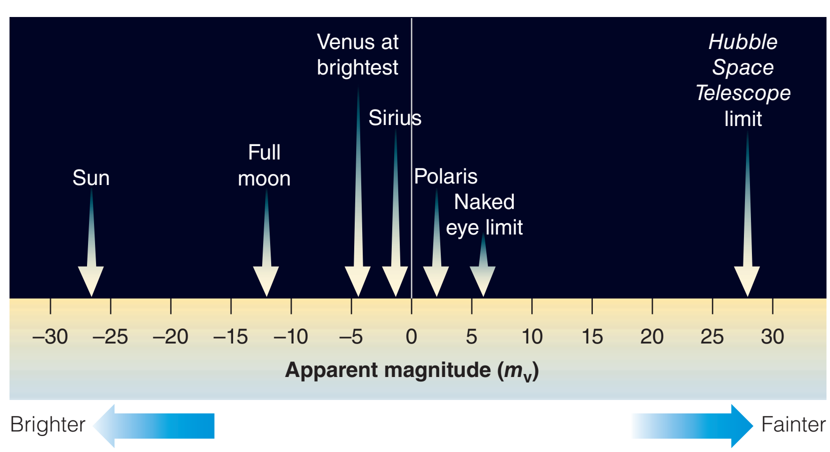
Almost 300 years later, Ptolemy used the magnitude system in his own catalog, and successive generations of astronomers have continued to use the system.

The ancient astronomers divided the stars into six classes. The brightest were called first-magnitude stars and those that were fainter, second-magnitude. The scale continued down to sixth-magnitude stars, the faintest visible to the human eye. Thus, the larger the magnitude number, the fainter the star. This makes sense if you think of the bright stars as first-class stars and the faintest visible stars as sixth-class stars.

Ancient astronomers could only estimate magnitudes, but modern astronomers can measure the brightness of stars to high precision, so they have made adjustments to the scale of magnitudes. Instead of saying that the star known by the charming name Chort (theta Leonis) is third magnitude, they can say its magnitude is 3.34. Accurate measurements show that some stars are actually brighter than magnitude 1.0. For example, Favorite Star Vega (alpha Lyrae) is so bright that its magnitude, 0.03, is almost zero. A few are so bright that the modern magnitude scale must extend into negative numbers (■ Figure 2-6). On this scale, Favorite Star Sirius, the brightest star in the sky, has a magnitude of −1.47. Modern astronomers have had to extend the faint end of the magnitude scale as well. The faintest stars you can see with your unaided eyes are about sixth magnitude, but if you use a telescope, you will see stars much fainter. Astronomers must use magnitude numbers larger than 6 to describe these faint stars.

These numbers are known as **apparent visual magnitudes (m_V)**, and they describe how the stars look to human eyes observing from Earth. Although some stars emit large amounts of infrared or ultraviolet light, human eyes can't see those types of radiation, and they are not included in the apparent visual magnitude. The subscript "V" stands for "visual" and reminds you that only visible light is included. Apparent visual magnitude also does not take into account the distance to the stars. Apparent visual magnitude ignores the effect of distance and tells you only how bright the star looks as seen from Earth.

Your interpretation of brightness is quite subjective, depending on both the physiology of human eyes and the psychology of perception. To be accurate you should refer to **flux**—the amount of light energy that hits one square meter in one second. This makes a precise definition of brightness. A simple relationship connects apparent visual magnitudes and flux (brightness) (**Reasoning with Numbers 2-1**). In this way, modern astronomers can measure the brightness of stars to high precision while still making comparisons to observations of apparent visual magnitude that go back to the time of Hipparchus.



■ **Figure 2-6**

The scale of apparent visual magnitudes extends into negative numbers to represent the brightest objects and to positive numbers larger than 6 to represent objects fainter than the human eye can see.

Reasoning with Numbers | 2-1

Magnitudes

Astronomers use a simple formula to convert between magnitudes and fluxes (brightness). If two stars have fluxes F_A and F_B, then the ratio of their fluxes is F_A/F_B. Modern astronomers have defined the magnitude scale so two stars that differ by five magnitudes have a flux ratio of exactly 100. Therefore, two stars that differ by one magnitude must have a flux ratio that equals the fifth root of 100, $\sqrt[5]{100}$ or $100^{0.2}$, which equals 2.51—that is, the light arriving at Earth from one star must be 2.51 times brighter than from the other. Two stars that differ by two magnitudes will have a flux ratio of 2.51×2.51, which is approximately 6.31, and so on (■ Table 2-1).

Example A: Suppose star C is third magnitude, and star D is ninth magnitude. What is their brightness ratio? **Solution:** The magnitude difference is six magnitudes, and Table 2-1 shows the corresponding flux ratio is 251. Therefore star C is 251 times brighter (has 251 times as much flux arriving at Earth) as star D.

A table is convenient, but for more precision you can express the relationship as a simple formula. The flux ratio F_A/F_B is equal to 2.51 raised to the power of the magnitude difference $m_B - m_A$:

$$\frac{F_A}{F_B} = (2.51)^{(m_B - m_A)}$$

Example B: If the magnitude difference is 6.32 magnitudes, what is the flux ratio? **Solution:** The flux ratio must be $2.51^{6.32}$. A pocket calculator tells you the answer: 336. The flux of light from star A is 336 times larger than the flux from star B. Note that, because of the way magnitudes are defined, star A has the larger flux and numerically smaller magnitude.

■ **Table 2-1 | Magnitude Differences and Flux Ratios**

Magnitude Difference	Corresponding Flux Ratio
0.00	1.00
1.00	2.51
2.00	6.31
3.00	15.8
4.00	39.8
5.00	100
6.00	251
7.00	631
8.00	1580
9.00	3980
10.00	10,000
:	:
:	:
15.00	1,000,000
20.00	100,000,000
25.0	10,000,000,000

On the other hand, when you know the flux ratio and want to find the magnitude difference, it is convenient to rearrange the previous formula and write it as:

$$m_B - m_A = 2.5 \log(F_A/F_B)$$

The expression *log* means logarithm to the base 10.

Example C: The light from Sirius has 24.2 times as much flux (is 24.2 times brighter than) light from Polaris. What is their magnitude difference? **Solution:** The magnitude difference is 2.5 times the logarithm of 24.2, which is written 2.5 log(24.2). Your calculator tells you the logarithm of 24.2 is 1.38, so the magnitude difference is 2.5×1.38, which equals 3.45 magnitudes.

2-2 The Sky and Its Motion

THE SKY ABOVE seems to be a great blue dome in the daytime and a sparkling ceiling at night. It was this ceiling that the first astronomers observed long ago as they tried to understand the night sky.

The Celestial Sphere

Ancient astronomers believed the sky was a great sphere surrounding Earth with the stars stuck on the inside like thumbtacks in a ceiling. Modern astronomers know that the stars are scattered through space at different distances, but it is still convenient to think of the sky as a great starry sphere enclosing Earth.

The Concept Art Spread **The Sky Around You** on pages 18–19 takes you on an illustrated tour of the sky. Throughout this book, these two-page art spreads introduce new concepts and new terms through photos and diagrams. These concepts and new terms are not discussed elsewhere, so examine the art spreads carefully. Notice that The Sky Around You introduces you to three important principles and 16 new terms that will help you understand the sky:

1. The sky appears to rotate westward around Earth each day, but that is a consequence of the eastward rotation of Earth. That rotation produces day and night. Notice how reference points on the *celestial sphere* such as the *zenith, nadir, horizon, celestial equator,* and the *north celestial pole* and *south*

How Do We Know? 2-1

Scientific Models

How can a scientific model be useful if it isn't entirely true? A scientific model is a carefully devised conception of how something works, a framework that helps scientists think about some aspect of nature, just as the celestial sphere helps astronomers think about the motions of the sky.

Chemists, for example, use colored balls to represent atoms and sticks to represent the bonds between them, kind of like Tinkertoys. Using these molecular models, chemists can see the three-dimensional shape of molecules and understand how the atoms interconnect. The molecular model of DNA proposed by Watson and Crick in 1953 led to our modern understanding of the mechanisms of genetics. You have probably seen elaborate ball-and-stick models of DNA, but does the molecule really look like Tinkertoys? No, but the model is both simple enough and accurate enough to help scientists think productively about molecules.

A scientific model is not a statement of truth; it does not have to be precisely true to be useful. In an idealized model, some complex aspects of nature can be simplified or omitted. The ball-and-stick model of a molecule doesn't show the relative strength of the chemical bonds, for instance. A model gives scientists a way to think about some aspect of nature but need not be true in every detail.

When you use a scientific model, it is important to remember the limitations of that model. If you begin to think of a model as true, it can be misleading instead of helpful. The celestial sphere, for instance, can help you think about the sky, but you must remember that it is only a model. The universe is much larger and much more interesting than this ancient scientific model of the heavens.

John Harwood/Photodisc/Jupiter Images

Balls represent atoms and rods represent chemical bonds in this model of a DNA molecule.

celestial pole define the four directions, *north point, south point, east point,* and *west point.*

2 Astronomers measure *angular distance* across the sky as angles and express them as degrees, *arc minutes,* and *arc seconds.* The same units are used to measure the *angular diameter* of an object.

3 What you can see of the sky depends on where you are on Earth. If you lived in Australia, you would see many constellations and asterisms invisible from North America, but you would never see the Big Dipper. How many *circumpolar constellations* you see depends on where you are. Remember your Favorite Star Alpha Centauri? It is in the southern sky and isn't visible from most of the United States. You could just glimpse it above the southern horizon if you were in Miami, Florida, but you could see it easily from Australia.

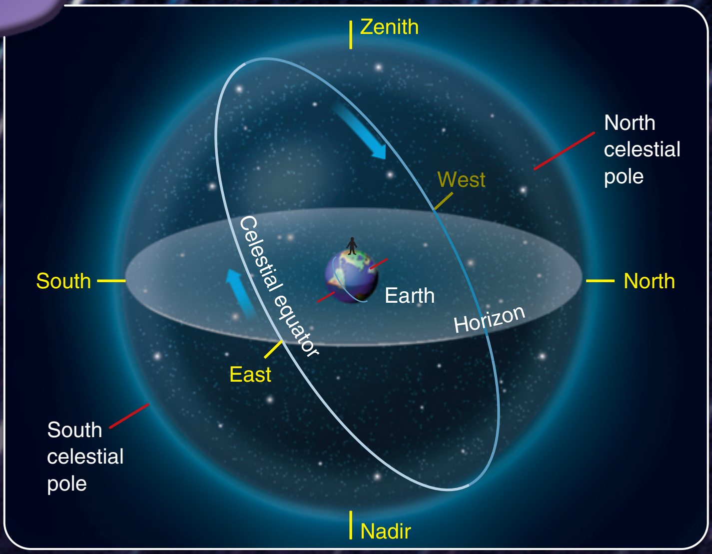
Pay special attention to the new terms on pages 18–19. You need to know these terms to describe the sky and its motions, but don't fall into the trap of just memorizing new terms. The goal of science is to understand nature, not to memorize definitions. Study the diagrams and see how the geometry of the celestial sphere and its motions produce the sky you see above you.

The celestial sphere is an example of a **scientific model**, a common feature of scientific thought (**How Do We Know? 2-1**). Notice that a scientific model does not have to be true to be useful. You will encounter many scientific models in the chapters that follow, and you will discover that some of the most useful models are highly simplified descriptions of the true facts.

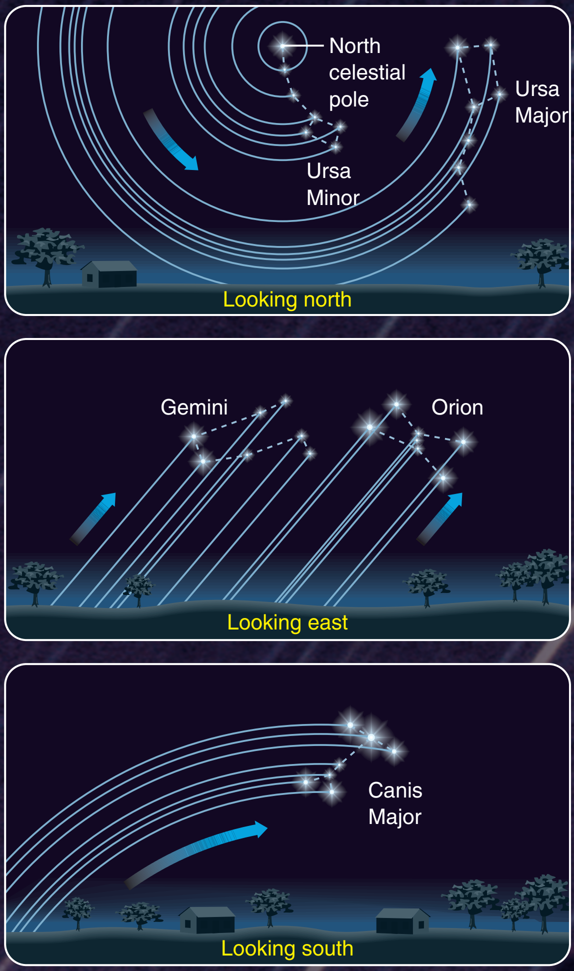
The Sky Around You

1 The eastward rotation of Earth causes the sun, moon, and stars to move westward in the sky as if the **celestial sphere** were rotating westward around Earth. From any location on Earth you see only half of the celestial sphere, the half above the **horizon.** The **zenith** marks the top of the sky above your head, and the **nadir** marks the bottom of the sky directly under your feet. The drawing at right shows the view for an observer in North America. An observer in South America would have a dramatically different horizon, zenith, and nadir.

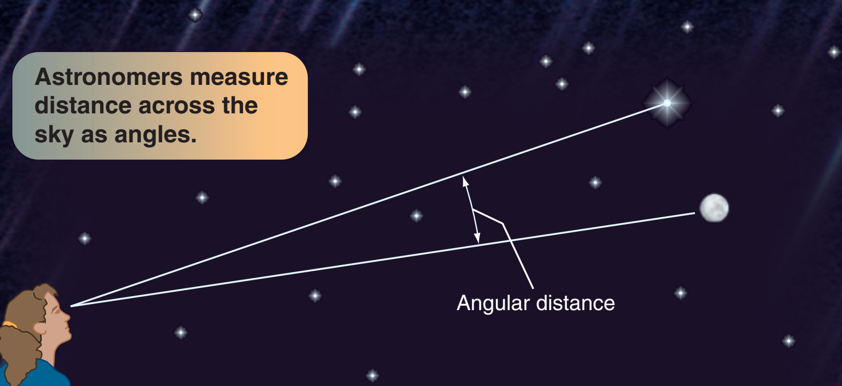
The apparent pivot points are the **north celestial pole** and the **south celestial pole** located directly above Earth's north and south poles. Halfway between the celestial poles lies the **celestial equator.** Earth's rotation defines the directions you use every day. The **north point** and **south point** are the points on the horizon closest to the celestial poles. The **east point** and the **west point** lie halfway between the north and south points. The celestial equator always meets the horizon at the east and west points.



AURA/NOAO/NSF

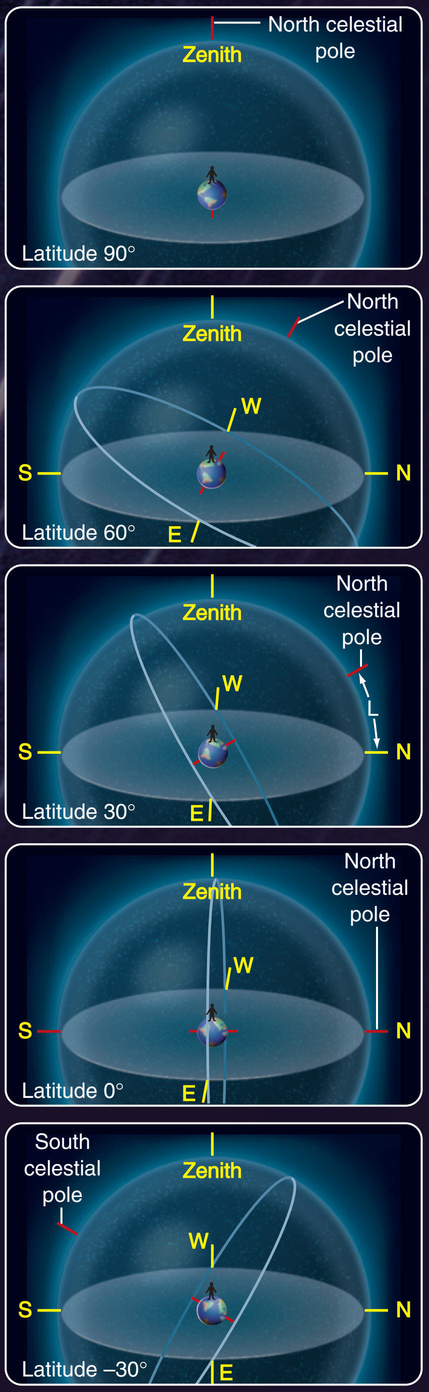


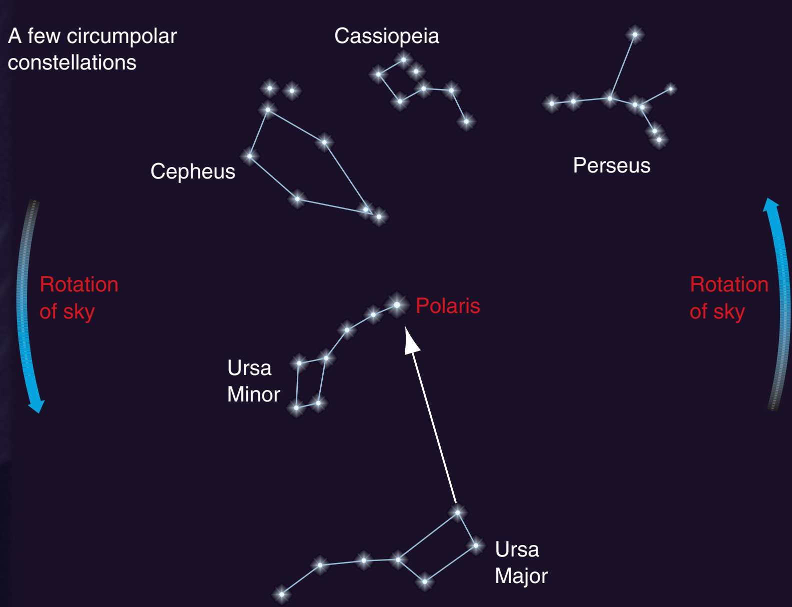
1a This time exposure of about 30 minutes shows stars as streaks, called star trails, rising behind an observatory dome. The camera was facing northeast to take this photo. The motion you see in the sky depends on which direction you look, as shown at right. Looking north, you see the Favorite Star Polaris, the North Star, located near the north celestial pole. As the sky appears to rotate westward, Polaris hardly moves, but other stars circle the celestial pole. Looking south from a location in North America, you can see stars circling the south celestial pole, which is invisible below the southern horizon.



2 Astronomers might say, "The star was two degrees from the moon." Of course, the stars are much farther away than the moon, but when you think of the celestial sphere, you can measure distance *on the sky* as an angle. The **angular distance** between two objects is the angle between two lines extending from your eye to the two objects. Astronomers measure angles in degrees, **arc minutes**, 1/60th of a degree, and **arc seconds**, 1/60th of an arc minute. Using the term "arc" avoids confusion with minutes and seconds of time. The **angular diameter** of an object is the angular distance from one edge to the other. The sun and moon are each about half a degree in diameter, and the bowl of the Big Dipper is about 10° wide.

3 What you see in the sky depends on your latitude as shown at right. Imagine that you begin a journey in the ice and snow at Earth's North Pole with the north celestial pole directly overhead. As you walk southward, the celestial pole moves toward the horizon, and you can see further into the southern sky. The angular distance from the horizon to the north celestial pole always equals your latitude (L)—the basis for celestial navigation. As you cross Earth's equator, the celestial equator would pass through your zenith, and the north celestial pole would sink below your northern horizon.





3a **Circumpolar constellations** are those that never rise or set. From mid-northern latitudes, as shown at left, you see a number of familiar constellations circling Polaris and never dipping below the horizon. As the sky rotates, the pointer stars at the front of the Big Dipper always point toward Polaris. Circumpolar constellations near the south celestial pole never rise as seen from mid-northern latitudes. From a high latitude such as Norway, you would have more circumpolar constellations, and from Quito, Ecuador, located on Earth's equator, you would have no circumpolar constellations at all.

This is a good time to eliminate a couple of **Common Misconceptions.** Lots of people, without thinking about it much, assume that the stars are not in the sky during the daytime. The stars are actually there day and night; they are just invisible during the day because the sky is lit up by sunlight. Also, many people insist that Favorite Star Polaris is the brightest star in the sky. It is actually the 51st visually brightest star. You now know that Polaris is important because of its position, not because of its brightness.

In addition to causing the obvious daily motion of the sky, Earth's rotation conceals a very slow celestial motion that can be detected only over centuries.

Precession

Over 2000 years ago, Hipparchus compared a few of his star positions with those recorded nearly two centuries earlier and realized that the celestial poles and equator were slowly moving across the sky. Later astronomers understood that this motion is caused by the toplike motion of Earth.

If you have ever played with a top or gyroscope, you have seen how the spinning mass resists any sudden change in the direction of its axis of rotation. The more massive the top and the more rapidly it spins, the more it resists your efforts to twist it out of position. But you probably recall that even the most rapidly spinning top slowly swings its axis around in a circle. The weight of the top tends to make it tip over, and this combines with its rapid rotation to make its axis sweep out the shape of a cone in a motion called **precession** (■ Figure 2-7a).

Earth spins like a giant top, but it does not spin upright in its orbit; it is tipped 23.4° from vertical. Earth's large mass and rapid rotation keep its axis of rotation pointed toward a spot near the star Polaris, and the axis would not wander if it were not for precession.

Because of its rotation, Earth has a slight bulge around its middle, and the gravity of the sun and moon pull on this bulge, tending to twist Earth's axis "upright" relative to its orbit. If Earth were a perfect sphere, it would not get twisted. Notice that gravity tends to make the top fall over, but it tends to twist Earth upright in its orbit. In both cases, the twisting of the axis of rotation combines with the rotation of the object and causes precession. Earth's axis precesses, taking about 26,000 years for one cycle (Figure 2-7b).

Because the locations of the celestial poles and equator are defined by Earth's rotational axis, precession slowly moves these reference marks. You would notice no change at all from night to night or year to year, but precise measurements can reveal the slow precession of the celestial poles and the resulting change in orientation of the celestial equator.

Over centuries, precession has dramatic effects. Egyptian records show that 4800 years ago the north celestial pole was near the star Thuban (alpha Draconis). The pole is now approaching Polaris and will be closest to it in about the year 2100. In about 12,000 years, the pole will have moved to within 5° of Vega (alpha Lyrae). Next time you glance at Favorite Star Vega, remind yourself that it will someday be a very impressive north star. Figure 2-7c shows the path followed by the north celestial pole. You will discover in later chapters that precession is common among rotating astronomical bodies.

What Are We? Along for the Ride

We humans are planet walkers. We live on the surface of a whirling planet, and as we look out into the depths of the universe we see the scattered stars near us. Because our planet spins, the stars appear to move westward across the sky in continuous procession.

The sky is a symbol of remoteness, order, and power, and that may be why so many cultures worship the sky in one way or another. Every culture divides the star patterns up to represent their heroes, gods, and symbolic creatures. Hercules looked down on the ancient Greeks, and the same stars represent the protector Båakkaataxpitchee (Bear Above) to the Crow people of North America. Among the hundreds of religions around the world, nearly all locate their gods and goddesses in the heavens. The gods watch over us from their remote and powerful thrones among the stars.

Our days are filled with necessary trivia, but astronomy enriches our lives by fitting us into the continuity of life on Earth. As you rush to an evening meeting, a glance at the sky will remind you that the sky carries our human heritage. Jesus, Moses, and Muhammad saw the same stars that you see. Aristotle watched the stars of Orion rise in the east and set in the west just as you do. Astronomy helps us understand what we are by linking us to the past of human experience on this planet.

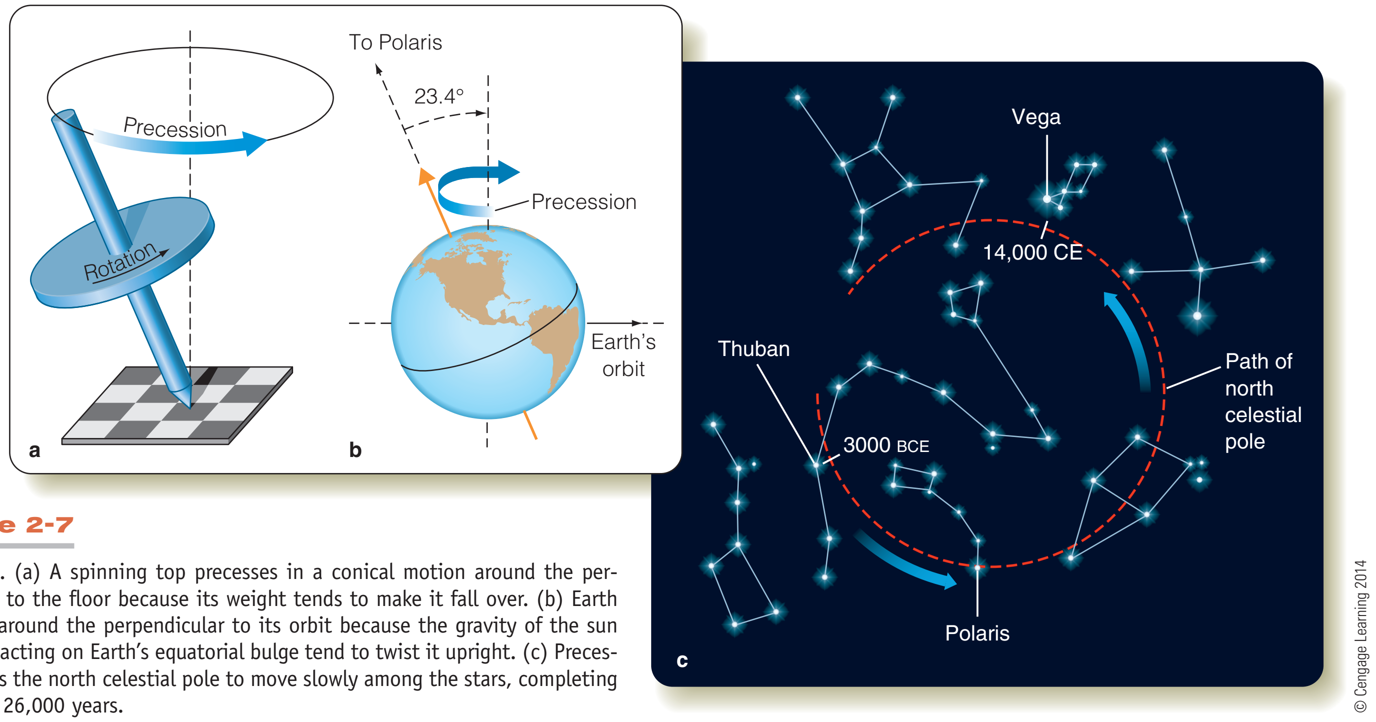


Figure 2-7

Precession. (a) A spinning top precesses in a conical motion around the perpendicular to the floor because its weight tends to make it fall over. (b) Earth precesses around the perpendicular to its orbit because the gravity of the sun and moon acting on Earth's equatorial bulge tend to twist it upright. (c) Precession causes the north celestial pole to move slowly among the stars, completing a circle in 26,000 years.

Study and Review

Summary

- Astronomers divide the sky into 88 **constellations (p. 12)** with modern boundaries defined by the International Astronomical Union (IAU). Although the constellations originated in Middle Eastern and Greek mythology, the names are Latin. Even constellations that are modern creations, added to fill in the spaces between the ancient figures, have Latin names. Named groups of stars that are not constellations are called **asterisms (p. 12).**
- The names of stars usually come from ancient Arabic, though modern astronomers often refer to a star by its constellation and a Greek letter assigned according to its brightness within the constellation.
- Astronomers refer to the brightness of stars using the **magnitude scale (p. 14).** First-magnitude stars are brighter than second-magnitude stars, which are brighter than third-magnitude stars, and so on. The magnitude you see when you look at a star in the sky is its **apparent visual magnitude,** m_v **(p. 15),** which includes only types of light visible to the human eye and does not take into account the star's distance from Earth.
- **Flux (p. 15)** is the amount of light energy that hits one square meter per second, a rigorous definition of brightness. The magnitude of a star is related mathematically to the flux of light received on Earth from the star, in other words, the star's brightness.
- The **celestial sphere (p. 18)** is a **scientific model (p. 17)** of the sky, to which the stars appear to be attached. Because Earth rotates eastward, the celestial sphere appears to rotate westward on its axis.
- The **north** and **south celestial poles (p. 18)** are the pivots on which the sky appears to rotate, and they define the four directions around the **horizon (p. 18):** the **north, south, east,** and **west points (p. 18).** The point directly overhead is the **zenith (p. 18),** and the point on the sky directly underfoot is the **nadir (p. 18).**
- The **celestial equator (p. 18),** an imaginary line around the sky above Earth's equator, divides the sky into northern and southern halves.
- Astronomers often refer to distances "on" the sky as if the stars, sun, moon, and planets were equivalent to spots painted on a plaster ceiling. These **angular distances (p. 19),** measured in degrees, **arc minutes (p. 19),** and **arc seconds (p. 19),** are unrelated to the true distance between the objects in light-years. The angular distance across an object is its **angular diameter (p. 19).**
- What you see of the celestial sphere depends on your latitude. Much of the Southern Hemisphere of the sky is not visible from northern latitudes. To see that part of the sky, you would have to travel southward over Earth's surface. **Circumpolar constellations (p. 19)** are those close enough to a celestial pole that they do not rise or set.
- The angular distance from the horizon to the north celestial pole always equals your latitude. This is the basis for celestial navigation.
- **Precession (p. 20)** is caused by the gravitational forces of the moon and sun acting on the equatorial bulge of the spinning Earth and causing its axis to sweep around in a conical motion like the motion of a top's axis. Earth's axis of rotation precesses with a period of 26,000 years, and consequently the celestial poles and celestial equator move slowly against the background of the stars.

Review Questions

1. Why have astronomers added modern constellations to the sky?
2. What is the difference between an asterism and a constellation? Give some examples.
3. What characteristic do stars in a constellation or asterism share?
4. Do people from other cultures on Earth see the same stars, constellations, and asterisms that you see?
5. How does the Greek-letter designation of a star give you a clue to its brightness?
6. How did the magnitude system originate in a classification of stars by brightness?
7. What does the word *apparent* mean in *apparent visual magnitude*?
8. In what ways is the celestial sphere a scientific model?
9. Why do astronomers use the word *on* to describe angles *on* the sky rather than angles *in* the sky?
10. If Earth did not rotate, could you define the celestial poles and celestial equator?
11. Where would you go on Earth if you wanted to be able to see both the north celestial pole and the south celestial pole at the same time?
12. Where would you go on Earth to place a celestial pole at your zenith?
13. Explain how to make a simple astronomical observation that would determine your latitude.
14. Why does the number of circumpolar constellations depend on the latitude of the observer?
15. How could you detect Earth's precession by examining star charts from ancient Egypt?
16. **How Do We Know?** How can a scientific model be useful if it isn't a correct description of nature?

Discussion Questions

1. All cultures on Earth named constellations. Why do you suppose this was such a common practice?
2. If you were lost at sea, you could find your approximate latitude by measuring the altitude of Polaris. But Polaris isn't exactly at the celestial pole. What else would you need to know to measure your latitude more accurately?

Problems

1. If light from one star is 40 times brighter (has 40 times more flux) than light from another star, what is their difference in magnitudes?
2. If two stars differ by 8.6 magnitudes, what is their flux ratio?
3. Star A has a magnitude of 2.5; Star B, 5.5; and Star C, 9.5. Which is brightest? Which are visible to the unaided eye? Which pair of stars has a flux ratio of 16?
4. By what factor is sunlight brighter than moonlight? (*Hint:* See Figure 2-6.)
5. If you are at a latitude of 35° north of Earth's equator, what is the angular distance from the northern horizon up to the north celestial pole? From the southern horizon down to the south celestial pole?

Learning to Look

1. Find the Southern Cross in the photo that opens this chapter.
2. This stamp shows the constellation Orion. Explain why this looks odd to residents of the Northern Hemisphere.

Great Debates

1. *Selling Space.* The Outer Space Treaty was signed by the United States, the United Kingdom, and the Soviet Union on October 10, 1967, and expressly forbids any government from claiming a celestial resource such as the moon (see United Nations Treaties and Principles on Space Law at www.oosa.unvienna.org/oosa/SpaceLaw/treaties.html). However, this does not prevent companies from dealing in certain transactions such as selling property on the moon, selling containers to ship your property to the moon, or selling the right to name stars or asteroids. Is the selling of these novelty gifts moral and/or ethical? If someone you know is considering purchasing one of these novelty gifts, how do you counsel them?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

2. *Labeling Stars.* Early cultures sometimes ordered star names by apparent brightness, but we now know the ordering is incorrect. For example, alpha Orionis is really dimmer than beta Orionis (see Figure 2-4). Should astronomers correct the original labeling of all stars so the labeling reflects the order of apparent brightness, or should the original labeling by the cultures remain?
   a. What's the evidence? Cite at least two cultures and their named constellations.
   b. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   c. Cite your sources.

3. *Clockwise/Counterclockwise Clocks.* On a sunny day, an obelisk (a tall, slender, four-sided monument such as the Washington Monument) casts a shadow on the ground that rotates clockwise in the Northern Hemisphere. The hour and minute hands on a clock also rotate in the clockwise direction. However, the sun's passage is in the opposite direction, or counterclockwise. Should clock hands be changed so that they rotate in the counterclockwise direction and number positions switched (that is, 1 becomes 11, 2 becomes 10, 3 becomes 9, and so on) to more accurately reflect the sun's passage?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

4. *A Curfew for Lights?* Should streetlights be dimmed at night to save on electric bills and hence your taxes? Some advantages would be that observatories would benefit from a curfew on bright lights, and you might be able to see more stars at night as well. If you are asked to sign a petition to send to your state government to enact a curfew on streetlight brightness, will you sign it?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

5. *Twinkle, Twinkle, Little Planet.* A little girl points to a bright nontwinkling object in the night sky, telling you the star is pretty. You know that stars twinkle, but planets do not. Do you correct the child?
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

Enhanced Web Assign

ENHANCED WebAssign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Celestial Sphere
- Rotation of the Sky

CengageNOW Virtual Astronomy Laboratories 2.0

Virtual ASTRONOMY LABORATORIES

An online, interactive way for you to explore astronomy, Virtual Astronomy Laboratories 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Laboratories. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

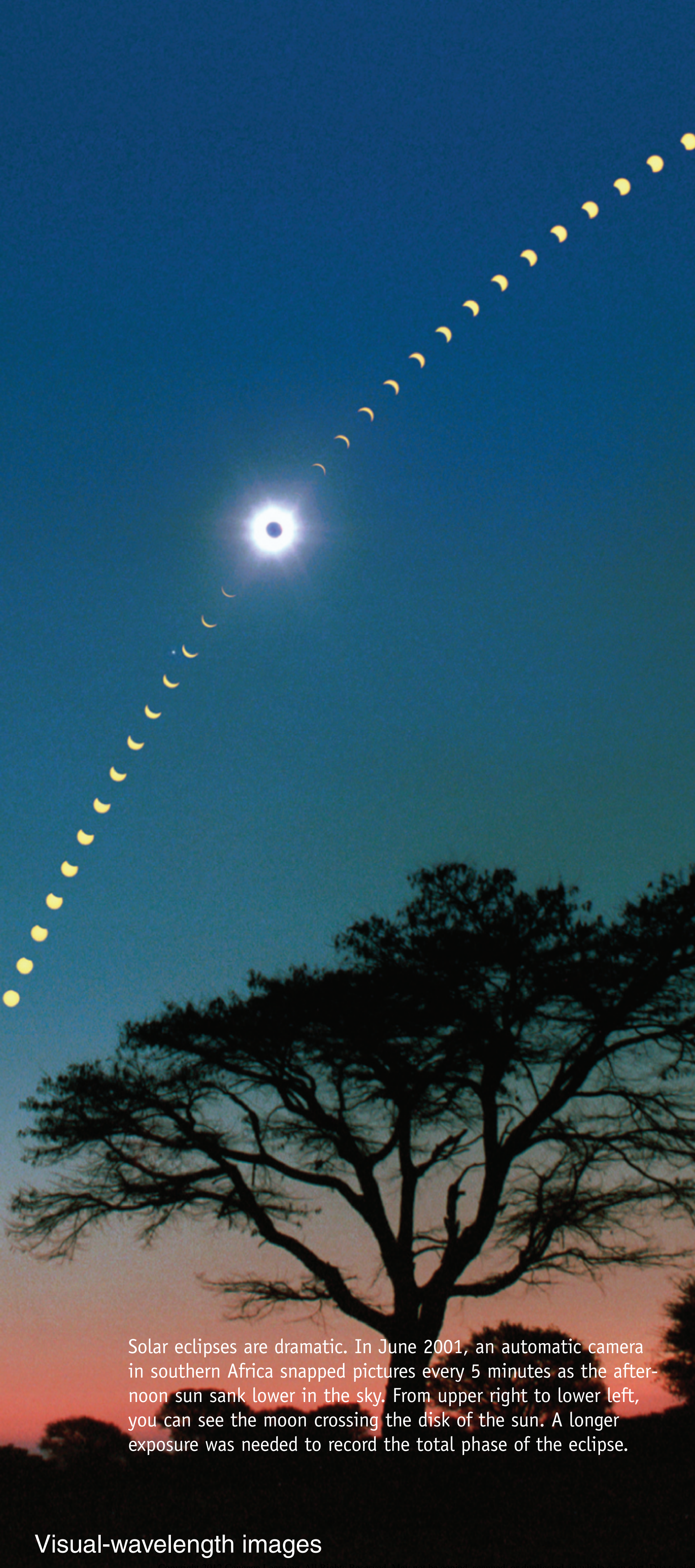
Virtual Astronomy Laboratory 2.1: Measurement and Unit Conversion

You see the rising full moon looming on the horizon, looking huge. But, if you happen to be camping in the mountains, a mountain peak on the horizon may appear bigger than the moon—you might be able to "cover" the moon with your thumb at arm's length but need your entire hand to block out the mountain. You know the mountain isn't bigger than the moon. What's going on? You need to make the distinction between how large something appears—say, whether you can cover it with your thumb or your hand—and how large it really is.

The true size (also known as the linear size) of an object is expressed in familiar units like miles or kilometers. The apparent size of an object—how large it seems to be, how much sky it takes up—is expressed in angular units. For example, the moon appears about 0.5 degrees in diameter. (Given that a full circle is 360 degrees, how many moons would it take, side-by-side, to stretch all the way around the horizon?) But then, how large is the moon, really? What is its linear diameter?

To find the true (linear) size of something, you need to know both its angular (apparent) size and its distance. As you read in the previous chapter, distances in astronomy are very difficult to determine. But, unless the distance of an object is known, its true size is unknown. A simple equation called the "small-angle formula" that is introduced in this chapter expresses the relationship between an object's apparent (angular) size, true (linear) size, and distance. If you know any two of those quantities you can calculate the third. The "small-angle" part of the name means that the formula is an approximation that is more accurate the smaller the angular size being considered.

Sections 3 and 4 of Virtual Astronomy Laboratory 1, "Measurement and Unit Conversion," help you practice using linear and angular units and employing the small-angle formula to calculate linear sizes, angular sizes, and distances. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.



Solar eclipses are dramatic. In June 2001, an automatic camera in southern Africa snapped pictures every 5 minutes as the afternoon sun sank lower in the sky. From upper right to lower left, you can see the moon crossing the disk of the sun. A longer exposure was needed to record the total phase of the eclipse.

2001 F. Espanak, www.MrEclipse.com

Visual-wavelength images

3

Cycles of the Sun and Moon

Guidepost

In the previous chapter you looked at the sky and saw how its apparent motion is caused by the daily rotation of Earth. In this chapter you will discover that the sun, moon, and planets move against the background of stars in cycles longer than one day. Some of those motions have direct influences on your life and produce dramatic sights in the sky. As you explore, you will find answers to four important questions:

- **What causes the seasons?**
- **How do astronomical cycles affect Earth's climate?**
- **Why does the moon go through phases?**
- **What causes lunar and solar eclipses?**

The cycles of the sky are elegant and dramatic, and you can understand them fully only if you know that Earth is a moving planet. Humans did not always know that. How we came to understand that Earth is a planet is the subject of the next chapter.

Even a man who is pure in heart
and says his prayers by night

May become a wolf when the wolfbane blooms
and the moon shines full and bright.

PROVERB FROM OLD WOLFMAN MOVIES

YOUR ALARM CLOCK and your calendar are astronomical instruments that track the motion of the sun in the sky. Furthermore, your calendar is divided into months, and that recognizes the monthly orbital motion of the moon. Your life is regulated by the cycles of the sky, and the most obvious cycle is that of the sun.

3-1 Cycles of the Sun

EARTH'S ROTATION ON ITS AXIS causes the cycle of day and night, but it is its motion around the sun in its orbit that defines the year. Notice an important distinction. **Rotation** is the turning of a body on its axis, but **revolution** means the motion of a body around a point outside the body. Consequently, astronomers are careful to say Earth rotates once a day on its axis and revolves once a year around the sun.

The Annual Motion of the Sun

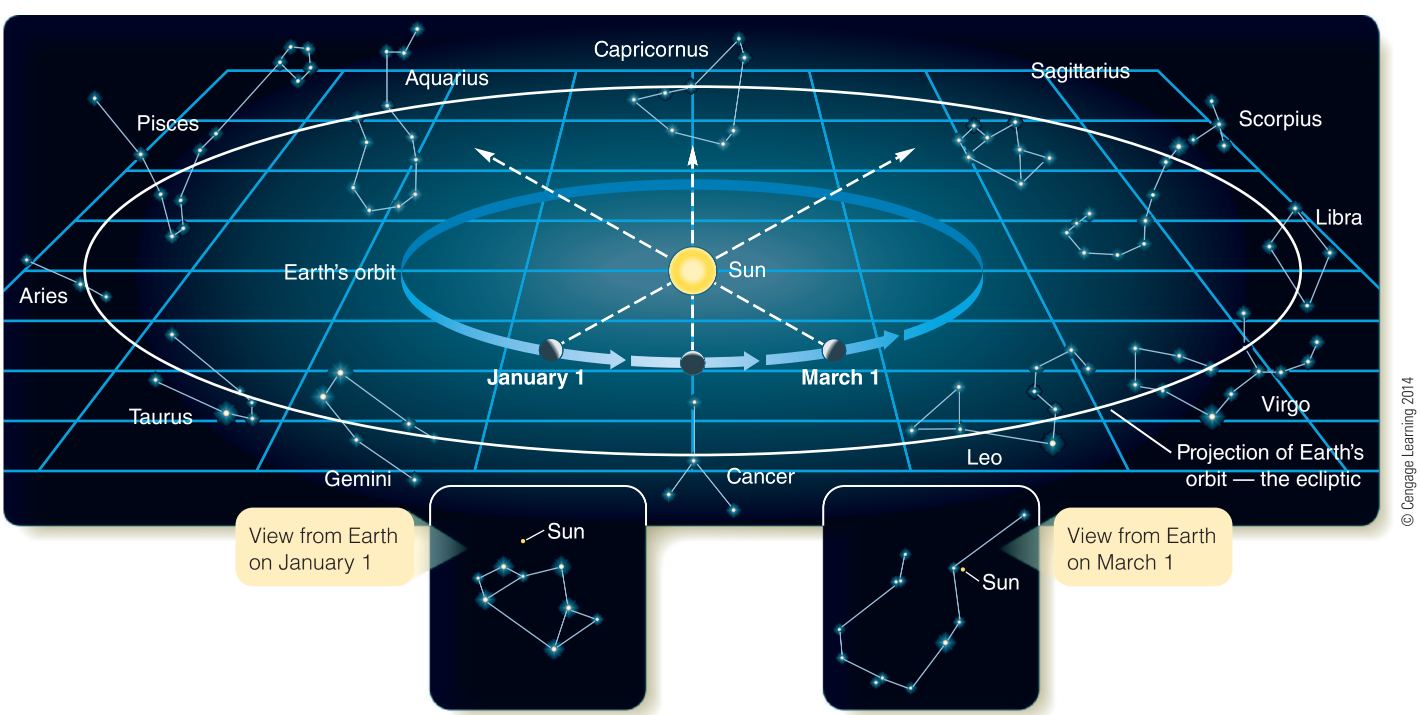
Even in the daytime, the sky is filled with stars, but the glare of sunlight fills Earth's atmosphere with scattered light, and you can see only the brilliant sun. If the sun were fainter, you would be able to see it rise in the morning in front of the stars. During the day, you would see the sun and the stars moving westward, and the sun would eventually set in front of the same stars. If you watched carefully as the day passed, you would notice that the sun was creeping slowly eastward against the background of stars. It would move a distance roughly equal to its own diameter between sunrise and sunset. This motion is caused by the motion of Earth in its nearly circular orbit around the sun.

For example, in January, you would see the sun in front of the constellation Sagittarius (■ Figure 3-1). As Earth moves along its circular orbit, the sun appears to move eastward among the stars. By March, you would see it in front of Aquarius.

Although people often say the sun is "in Sagittarius" or "in Aquarius," it isn't really correct to say the sun is "in" a constellation. The sun is only 1 AU away, and most of the stars visible in the sky are millions of times more distant. Nevertheless, in March

■ **Figure 3-1**

Earth's orbit is a nearly perfect circle, but it is shown in an inclined view in this diagram and consequently looks oval. Earth's motion around the sun makes the sun appear to move against the background of the stars. Earth's orbit is thus projected on the sky as the circular path of the sun, the ecliptic. If you could see the stars in the daytime, you would notice the sun crossing in front of the distant constellations as Earth moves along its orbit.



of each year, the sun crosses in front of the stars that make up Aquarius, and people use the common expression, "The sun is in Aquarius."

The apparent path of the sun against the background of stars is called the **ecliptic**. If the sky were a great screen, the ecliptic would be the shadow cast by Earth's orbit. That is why the ecliptic is often called the projection of Earth's orbit on the sky.

Earth circles the sun in 365.26 days, and consequently the sun appears to circle the sky in the same period. That means the sun, traveling 360° around the ecliptic in 365.26 days, travels about 1° eastward in 24 hours, about twice its angular diameter. You don't notice this apparent motion of the sun because you can't see the stars in the daytime, but it does have an important consequence that you do notice—the seasons.

The Seasons

Earth would not experience seasons if it rotated upright in its orbit, but it does have seasons because its axis of rotation is tipped 23.4° from the perpendicular to its orbit. Another way to say this is that Earth's equator is inclined 23.4° to its orbit. Study **The Cycle of the Seasons** on pages 26–27 and notice two important principles and six new terms:

1. Because Earth's axis of rotation is inclined 23.4°, the sun moves into the northern sky in the spring and into the southern sky in the fall. That causes the cycle of the seasons. Notice how the *vernal equinox,* the *summer solstice,* the *autumnal equinox,* and the *winter solstice* mark the beginning of the seasons. Earth's elliptical orbit is very nearly circular, and as it travels from *perihelion* to *aphelion,* its distance from the sun varies only slightly and is not the cause of the seasons.

2. Earth goes through a cycle of seasons because of changes in the amount of solar energy that Earth's northern and southern hemispheres receive at different times of the year. Because of circulation patterns in Earth's atmosphere, the northern and southern hemispheres are mostly isolated from each other and exchange little heat. When one hemisphere receives more solar energy than the other, it grows rapidly warmer.

Notice that the seasons in Earth's southern hemisphere are reversed with respect to those in the northern hemisphere, Australia and other lands in the southern hemisphere experience winter from June 21 to September 22, and summer from December 21 to March 20.

Now you can set your friends straight if they mention two of the most **Common Misconceptions** about the seasons. First, the seasons don't occur because Earth moves closer to or farther from the sun. If that were the cause, both of Earth's hemispheres would experience winter at the same time, and that's not what happens. Earth's orbit is nearly circular. Its ditance from the sun varies by only a few percent from winter to summer, and that isn't enough to cause the seasons. Second, it is not easier to stand a raw egg on end on the day of the vernal equinox! Have you heard that one? Radio and TV personalities love to talk about it, but it just isn't true. It is one of the silliest misconceptions in science. You can stand a raw egg on end any day of the year if you have steady hands. (*Hint:* It helps to shake the egg really hard to break the yolk inside so it can settle to the bottom.)

Motions of the Planets

The planets of our solar system produce no visible light of their own; they are visible only by reflected sunlight. Mercury, Venus, Mars, Jupiter, and Saturn are all easily visible to the unaided eye and look like stars, but Uranus is usually too faint to be seen, and Neptune is never bright enough.

All of the planets of the solar system, including Earth, move in nearly circular orbits around the sun. If you were looking down on the solar system from the north celestial pole, you would see the planets moving in the same counterclockwise direction around their orbits, with the planets farthest from the sun moving the slowest. Seen from Earth, the outer planets move slowly eastward* along the ecliptic. In fact, the word *planet* comes from the Greek word meaning "wanderer." Mars moves completely around the ecliptic in slightly less than 2 years, but Saturn, being farther from the sun, takes nearly 30 years.

Mercury and Venus also stay near the ecliptic, but they move differently from the other planets. They have orbits inside Earth's orbit, and that means they are never seen far from the sun in the sky. Observed from Earth, they move eastward away from the sun and then back toward the sun, crossing the near part of their orbit. They continue moving westward away from the sun and then move back, crossing the far part of their orbit before they move out east of the sun again. To find one of these planets, you need to look above the western horizon just after sunset or above the eastern horizon just before sunrise. Venus is easier to locate because it is brighter and because its larger orbit carries it higher above the horizon than does Mercury's (■ Figure 3-2). Mercury's orbit is so small that it can never get farther than 28° from the sun. Consequently, it is hard to see against the sun's glare and is often hidden in the clouds and haze near the horizon.

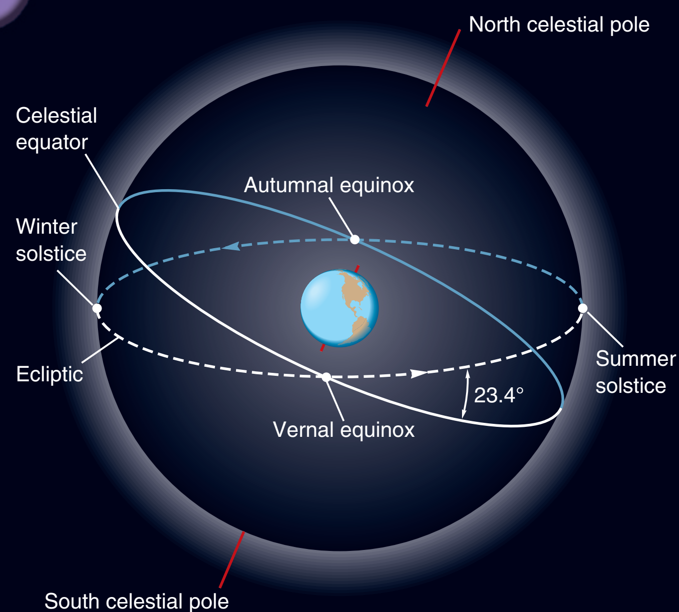
By tradition, any planet visible in the evening sky is called an **evening star**, even though planets are not stars. Similarly, any planet visible in the sky shortly before sunrise is called a **morning star**. Perhaps the most beautiful is Venus, which can become as bright as magnitude −4.7. As Venus moves around its orbit, it can dominate the western sky each evening for many weeks, but eventually its orbit carries it back toward the sun, and it is lost in

*You will discover occasional exceptions to this eastward motion in Chapter 4.

The Cycle of the Seasons

1 You can use the celestial sphere to help you think about the seasons. The celestial equator is the projection of Earth's equator on the sky, and the ecliptic is the projection of Earth's orbit on the sky. Because Earth is tipped in its orbit, the ecliptic and equator are inclined to each other by 23.4° as shown at right. As the sun moves eastward around the sky, it spends half the year in the southern half of the sky and half the year in the northern half. That causes the seasons.

The sun crosses the celestial equator going northward at the point called the **vernal equinox.** The sun is at its farthest north at the point called the **summer solstice.** It crosses the celestial equator going southward at the **autumnal equinox** and reaches its most southern point at the **winter solstice.**



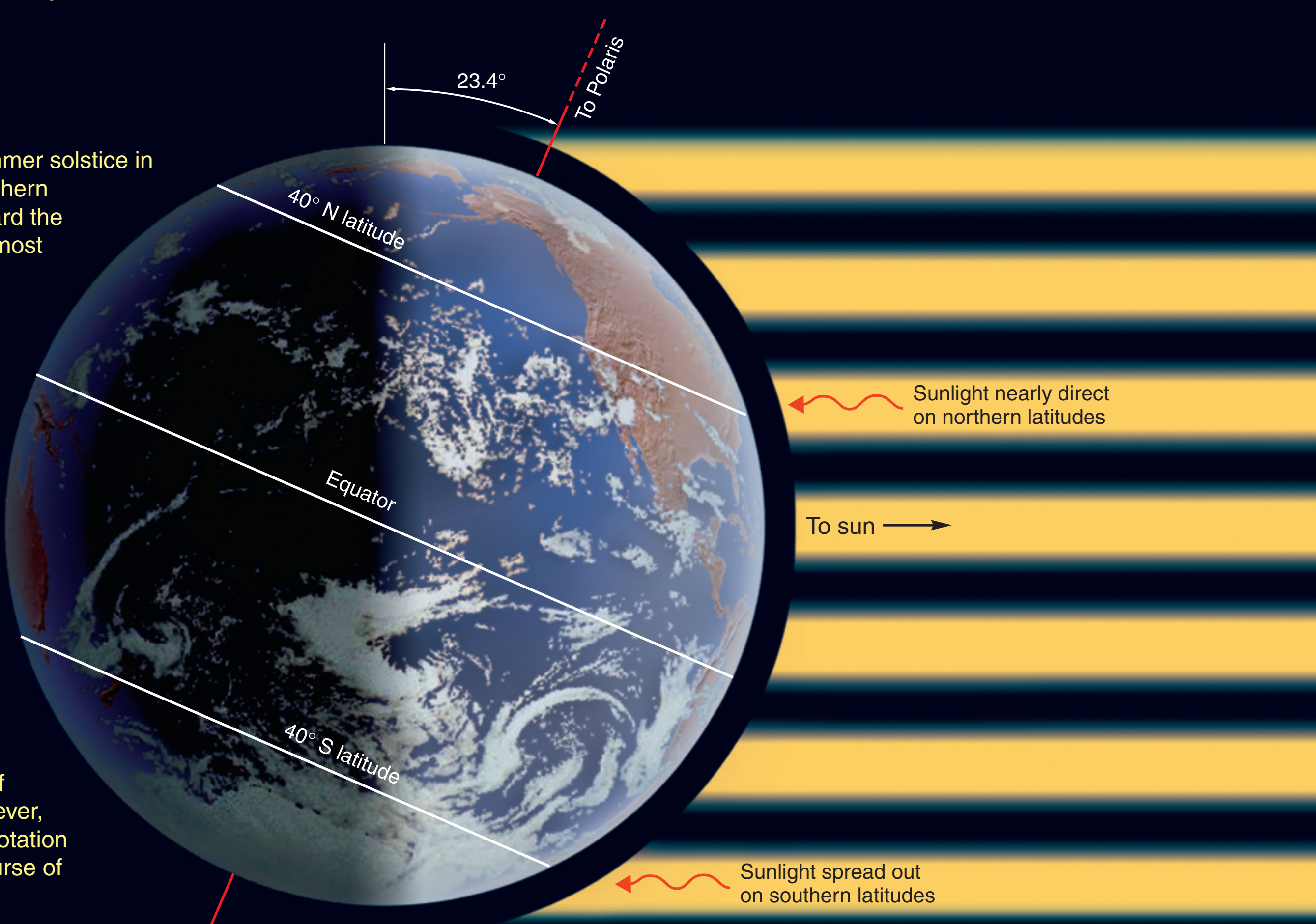
1a The seasons are defined by the dates when the sun crosses these four points, as shown in the table at the right. *Equinox* comes from the word for "equal"; the day of an equinox has equal amounts of daylight and darkness. *Solstice* comes from the words meaning "sun" and "stationary." *Vernal* comes from the word for "green." The "green" equinox marks the beginning of spring in the northern hemisphere.

Event	Date*	N. Hemisphere
Vernal equinox	March 20	Spring begins
Summer solstice	June 22	Summer begins
Autumnal equinox	September 22	Autumn begins
Winter solstice	December 21	Winter begins

* Give or take a day due to leap year and other effects.

1b On the day of the summer solstice in late June, Earth's northern hemisphere is inclined toward the sun, and sunlight shines almost straight down at northern latitudes. At southern latitudes, sunlight strikes the ground at an angle and spreads out. North America has warm weather, and South America has cool weather.

Earth's axis of rotation points toward Polaris, and, like a top, the spinning Earth holds its axis fixed as it orbits the sun. On one side of the sun, Earth's northern hemisphere leans toward the sun; on the other side of its orbit, it leans away. However, the direction of the axis of rotation does not change in the course of a year.



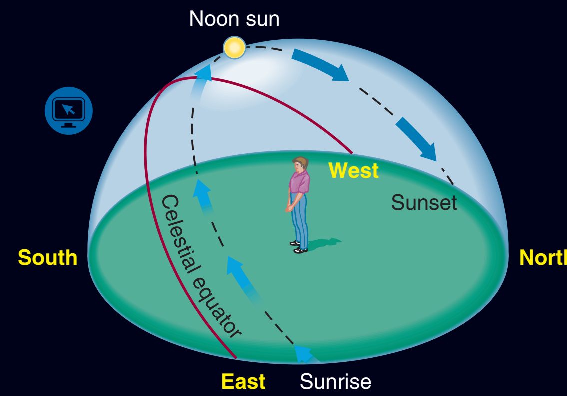
NASA

Earth at summer solstice

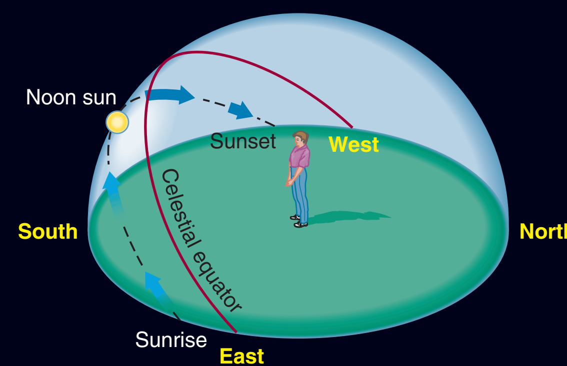
1c Light striking the ground at a steep angle spreads out less than light striking the ground at a shallow angle. Light from the summer solstice sun strikes northern latitudes from nearly overhead and is concentrated.

Light from the winter solstice sun strikes northern latitudes at a much steeper angle and spreads out. The same amount of energy is spread over a larger area, so the ground receives less energy from the winter sun.

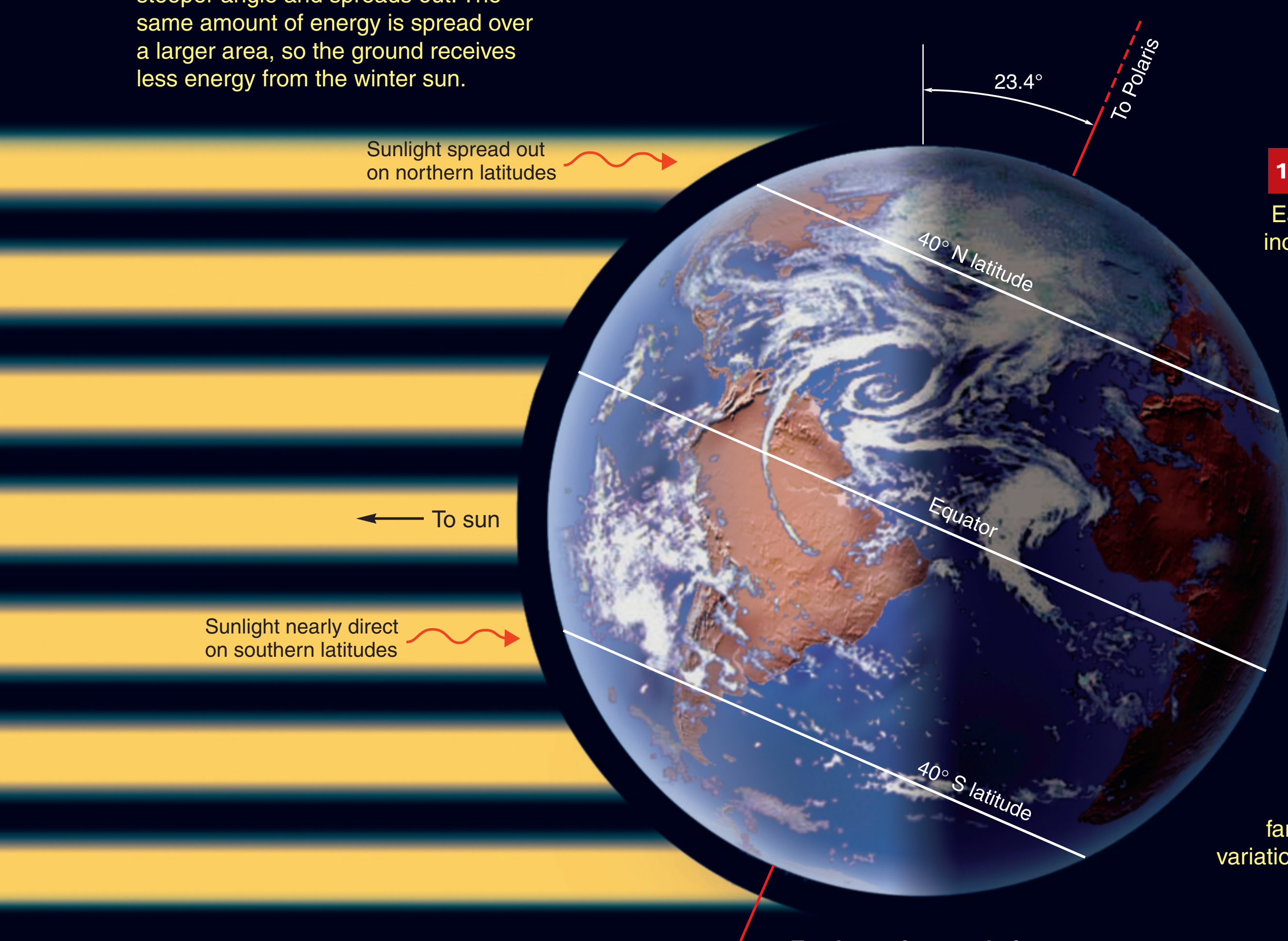
2 The two causes of the seasons are shown at right for someone in the northern hemisphere. First, the noon summer sun is higher in the sky and the winter sun is lower, as shown by the longer winter shadows. Thus winter sunlight is more spread out. Second, the summer sun rises in the northeast and sets in the northwest, spending more than 12 hours in the sky. The winter sun rises in the southeast and sets in the southwest, spending less than 12 hours in the sky. Both of these effects mean that northern latitudes receive more energy from the summer sun, and summer days are warmer than winter days.



At summer solstice



At winter solstice



Earth at winter solstice

NASA

1d On the day of the winter solstice in late December, Earth's northern hemisphere is inclined away from the sun, and sunlight strikes the ground at an angle and spreads out. At southern latitudes, sunlight shines almost straight down and does not spread out. North America has cool weather and South America has warm weather.

Earth's orbit is only very slightly elliptical. About January 3, Earth is at **perihelion,** its closest point to the sun, when it is only 1.7 percent closer than average. About July 5, Earth is at **aphelion,** its most distant point from the sun, when it is only 1.7 percent farther than average. This small variation does not significantly affect the seasons.

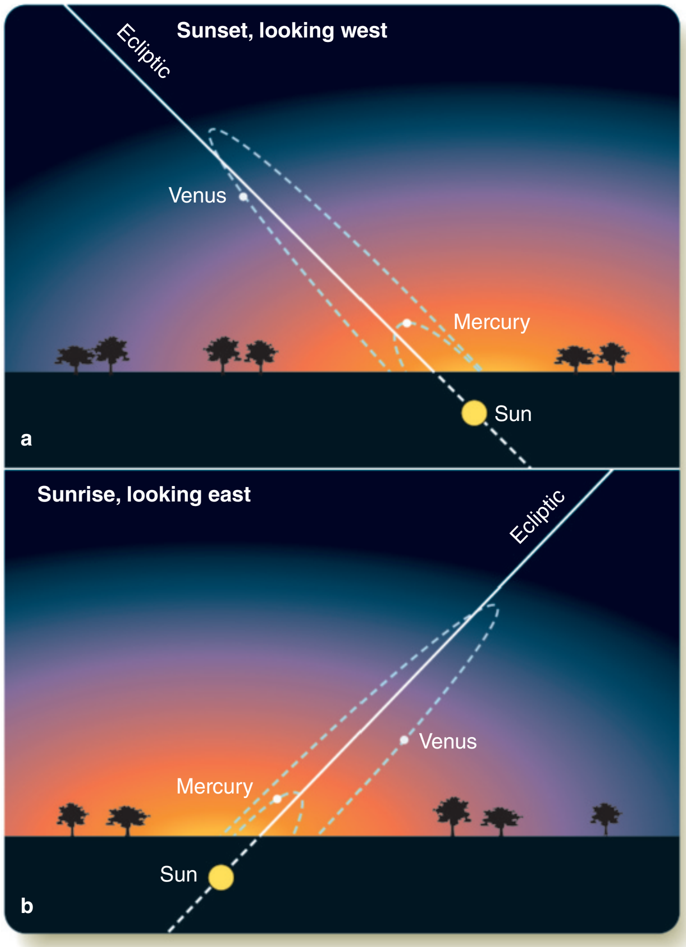


Figure 3-2

Mercury and Venus follow orbits that keep them near the sun, and they are visible only soon after sunset or before sunrise when the brilliance of the sun is hidden below the horizon. Venus takes 584 days to move from the morning sky to the evening sky and back again, but Mercury zips around in only 116 days.

the haze near the horizon. In a few weeks, it reappears in the dawn sky, a brilliant morning star.

The cycles of the sky are so impressive that it is not surprising that people have strong feelings about them. Ancient peoples saw the motion of the sun around the ecliptic as a powerful influence on their daily lives, and the motion of the planets along the ecliptic seemed similarly meaningful. The ancient superstition of astrology is based on the cycles of the sun and planets around the sky. You have probably heard of the **zodiac**, a band around the sky extending 9° above and below the ecliptic. The signs of the zodiac take their names from the 12 principal constellations along the ecliptic. A **horoscope** is just a diagram showing the location of the sun, moon, and planets around the ecliptic and their position above or below the horizon for a given date and time. Centuries ago, astrology was an important part of astronomy, but the two are now almost exact opposites—astronomy is a science that depends on evidence, and astrology is a superstition that survives in spite of evidence (**How Do We Know? 3-1**). The signs of the zodiac are no longer important in astronomy.

3-2 Astronomical Influences on Earth's Climate

The seasons are produced by the annual motion of Earth around the sun, but subtle changes in that motion can have dramatic affects on climate. You don't notice these changes during your lifetime; but, over thousands of years, they can bury continents under glaciers.

Earth has gone through ice ages when the worldwide climate was cooler and dryer and thick layers of ice covered polar regions and extended partway to the equator. Between ice ages, Earth is warmer and there are no ice sheets even at the poles. Scientists have found evidence of at least four ice ages in Earth's past. One occurred 2.5 billion years ago, but the other three that have been identified have all occurred in the last billion years. There were probably others, but evidence of early ice ages is usually erased by more recent ice sheets. The lengths of ice ages range from a few tens of millions of years to a few hundred million years. The most recent ice age began only about 3 million years ago and is still going on. You are living during one of the periodic episodes during an ice age when Earth grows slightly warmer and the glaciers melt back and do not extend as far from the poles. The current warm period began about 12,000 years ago.

Ice ages seem to occur with a period of roughly 250 to 300 million years, and glaciers advance and melt back during these ice ages in a complicated pattern that involves cycles of 40,000 to 100,000 years. (These cycles have no connection with global warming, which can produce changes in Earth's climate over just a few decades. Global warming is discussed in Chapter 9.) Evidence shows that these slow cycles of the ice ages have an astronomical origin.

The Hypothesis

Sometimes a theory or hypothesis is proposed long before scientists can find the critical evidence to test it. That happened in 1920 when Yugoslavian meteorologist Milutin Milankovitch proposed what became known as the **Milankovitch hypothesis**—that small changes in Earth's orbit, precession, and inclination affect Earth's climate and can cause ice ages. You should examine each of these motions separately.

First, Earth's elliptical orbit differs only very slightly from a circle, but astronomers know that because of gravitational

How Do We Know? 3-1

Pseudoscience

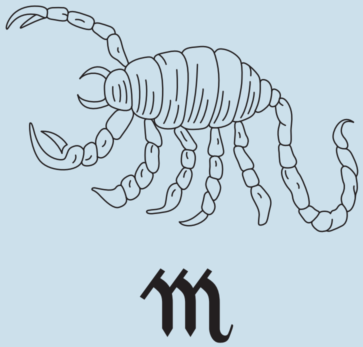
What is the difference between a science and a pseudoscience? Astronomers have a low opinion of beliefs such as astrology, not only because they are groundless but also because they *pretend* to be a science. They are **pseudosciences,** from the Greek *pseudo,* meaning "false."

A pseudoscience is a set of beliefs that appear to be based on scientific ideas but that fail to obey the most basic rules of science. For example, in the 1970s a claim was made that pyramidal shapes focus cosmic forces on anything underneath and might even have healing properties. Believers claimed that a pyramid made of paper, plastic, or other materials would preserve fruit, sharpen razor blades, and do other miraculous things. Many books promoted the idea of the special power of pyramids, and this idea led to a popular fad.

A key characteristic of science is that its claims can be tested and verified. In this case, simple experiments showed that any shape, not just a pyramid, protects a piece of fruit from airborne spores and allows it to dry without rotting. Likewise, any shape allows oxidation to improve the cutting edge of a razor blade. Because experimental evidence contradicted the claim and because supporters of the theory declined to abandon or revise their claims, you can recognize pyramid power as a pseudoscience. Disregard of contradictory evidence and alternate explanations is a sure sign of a pseudoscience.

Pseudoscientific claims can be self-fulfilling. For example, some believers in pyramid power slept under pyramidal tents to improve their rest. There is no logical mechanism by which such a tent could affect a sleeper, but because people wanted and expected the claim to be true, they reported that they slept more soundly. Vague claims based on personal testimony that cannot be tested are another sign of a pseudoscience.

Astrology is a pseudoscience. It has been tested over and over for centuries, and it doesn't work. Nevertheless, many people believe in astrology despite contradictory evidence. Many pseudosciences appeal to our need to understand and control the world around us. Some such claims involve medical cures, ranging from using magnetic bracelets and crystals to focus mystical power to astonishingly expensive, illegal, and dangerous treatments for cancer. Logic is a stranger to pseudoscience, but human fears and needs are not.



Astrology may be the oldest pseudoscience.

interactions with the other planets, the elliptical shape varies slightly over a period of about 100,000 years. Earth's orbit at present carries it 1.7 percent closer than average to the sun during Northern Hemisphere winters and 1.7 percent farther away in Northern Hemisphere summers. This makes northern summers very slightly cooler, and that is critical—most of the landmass where ice can accumulate is in the Northern Hemisphere. When Earth's orbit becomes more elliptical, northern summers might be too cool to melt all of the snow and ice from the previous winter. That would make glaciers grow larger.

A second factor is also at work. Precession causes Earth's axis to sweep around a cone over a period of about 26,000 years, and that gradually changes the points in Earth's orbit where a given hemisphere experiences the seasons. Northern Hemisphere summers now occur when Earth is 1.7 percent farther from the sun, but in 13,000 years northern summers will occur on the other side of Earth's orbit where Earth is 1.7 percent closer to the sun. Northern summers will be warmer, which could melt all of the previous winter's snow and ice and prevent the growth of glaciers.

The third factor is the inclination of Earth's equator to its orbit, currently at 23.4°. Because of gravitational tugs of the moon, sun, and even the other planets, this angle varies from 22° to 24° over a period of roughly 41,000 years. When the inclination is greater, seasons are more severe.

In 1920 Milankovitch proposed that these three factors cycle against each other to produce complex periodic variations in Earth's climate and the advance and retreat of glaciers (■ Figure 3-3a). Unfortunately, no evidence was available to test the theory in 1920, and scientists treated it with skepticism. Many thought it was laughable.

The Evidence

By the middle 1970s, Earth scientists were able to collect the data that Milankovitch had lacked. Oceanographers drilled deep into the seafloor to collect long cores of sediment. In the laboratory, geologists could take samples from different depths in the cores and determine the age of the samples and the temperature of the oceans when they were deposited on the seafloor. From this, scientists constructed a history of ocean temperatures that convincingly matched the predictions of the Milankovitch hypothesis (Figure 3-3b).

The evidence seemed very strong, and by the 1980s the Milankovitch hypothesis was widely considered the leading

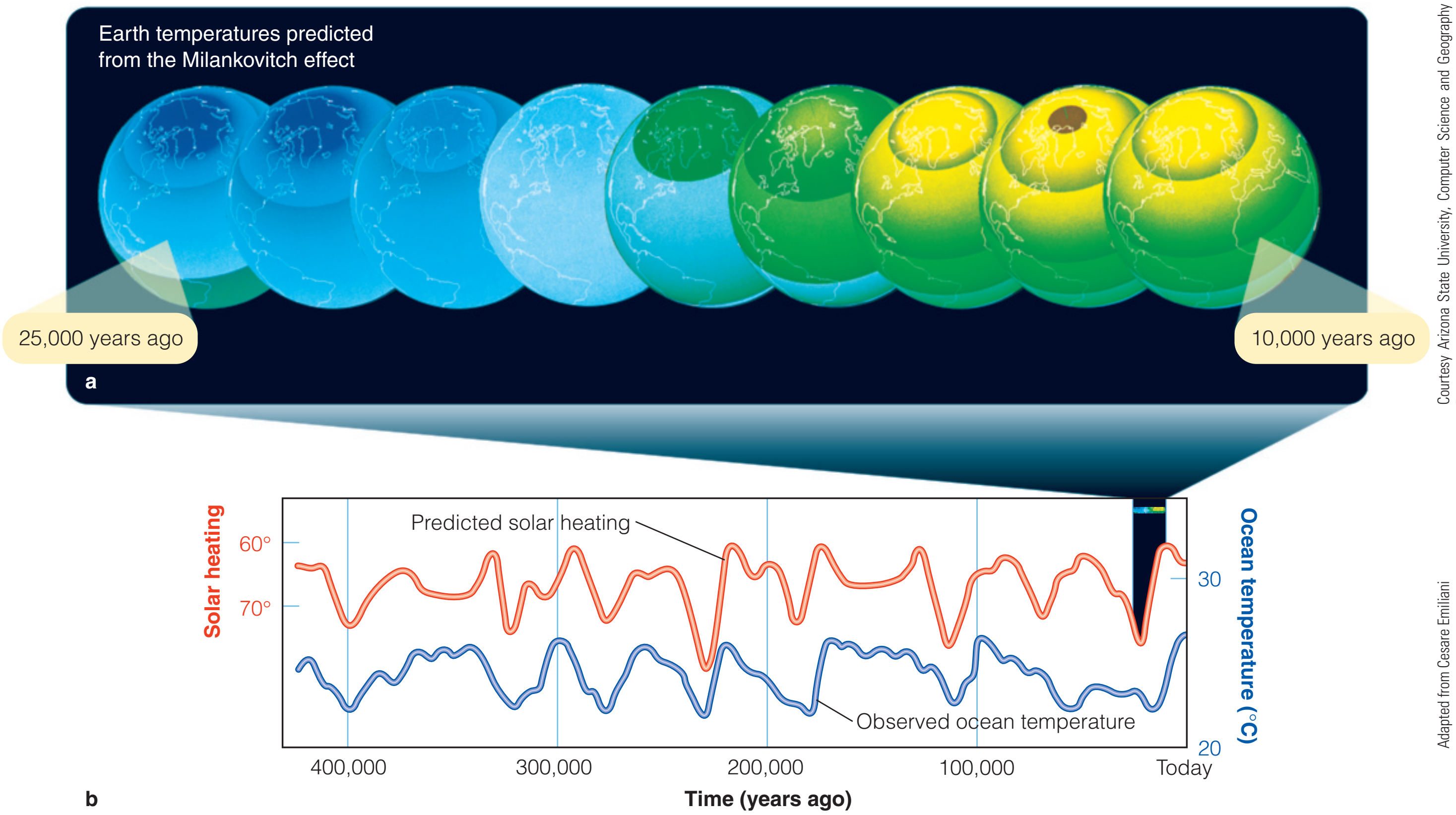


Figure 3-3

(a) Calculations based on the Milankovitch effect can predict temperature over time. The warming illustrated by the globes shown here took place from 25,000 to 10,000 years ago and ended the last glaciation. (b) Over the last 400,000 years, changes in ocean temperatures measured from fossils found in sediment layers from the seabed match changes in solar heating predicted by the Milankovitch effect.

hypothesis. But science follows a mostly unstated set of rules that holds that a hypothesis must be tested over and over against all available evidence (**How Do We Know? 3-2**). In 1988 scientists discovered contradictory evidence.

For 500,000 years rainwater has collected in a deep crack in Nevada called Devil's Hole. That water has deposited the mineral calcite in layer on layer on the walls of the crack. It isn't easy to get to, and scientists had to dive with scuba gear to drill out samples of the calcite, but it was worth the effort. Back in the laboratory, they could determine the age of each layer in their core samples and the temperature of the rainwater that had formed the calcite in each layer. That gave them a history of temperatures at Devil's Hole that spanned many thousands of years, and the results were a surprise. The evidence seemed to show that Earth had begun warming up thousands of years too early for the last ice age to have been caused by the Milankovitch cycles.

These contradictory findings are irritating because we humans naturally prefer certainty, and it isn't unusual for people to ignore evidence that does not support their beliefs. Such circumstances are common in science, and scientists must learn to weigh such contradictory evidence with care. The disagreement between the ocean floor samples and the Devil's Hole samples triggered a scramble to understand the problem. Were the age determinations of one or the other set of samples wrong? Were the calculations of ancient temperatures wrong? Or were scientists misunderstanding the significance of the evidence?

In 1997 a new study of the ages of the samples confirmed that those from the ocean floor are correctly dated. But the same study found that the ages of the Devil's Hole samples are also correct. Evidently the temperatures at Devil's Hole record local climate changes in the region that became the southwestern United States. The ocean floor samples record global

How Do We Know? | 3-2

Evidence as the Foundation of Science

Why is evidence critical in science? From colliding galaxies to the inner workings of atoms, scientists love to speculate and devise theories, but all scientific knowledge is ultimately based on evidence from observations and experiments. Evidence is reality, and scientists constantly check their ideas against reality.

When you think of evidence, you probably think of criminal investigations in which detectives collect fingerprints and eyewitness accounts. In court, that evidence is used to try to understand the crime, but there is a key difference in how lawyers and scientists use evidence. A defense attorney can call a witness and intentionally fail to ask a question that would reveal evidence harmful to the defendant. In contrast, the scientist must be objective and not ignore any known evidence. The attorney is presenting only one side of the case, but the scientist is searching for the truth. In a sense, the scientist must deal with the evidence as both the prosecution and the defense.

It is a characteristic of scientific knowledge that it is supported by evidence. A scientific statement is more than an opinion or a speculation because it has been tested objectively against reality.

As you read about any science, look for the evidence in the form of observations and experiments. Every theory or conclusion should have supporting evidence. If you can find and understand the evidence, the science will make sense. All scientists, from astronomers to zoologists, demand evidence. You should, too.

Dorling Kindersley/Getty Images

Fingerprints are evidence of past events.

climate changes, and they fit well with the Milankovitch hypothesis. This has given scientists renewed confidence in the Milankovitch hypothesis, and although it is widely accepted today, it is still being tested whenever scientists can find more evidence.

As you review this section, notice that it is a **scientific argument**, a careful presentation of theory and evidence in a logical discussion. **How Do We Know? 3-3** expands on the ways scientists organize their ideas in logical arguments. Throughout this book, many chapter sections end with short reviews entitled "Scientific Argument." These feature a review question, which is then analyzed in a scientific argument. A second question gives you a chance to build your own scientific argument on a related topic. You can use these "Scientific Argument" features not only to review chapter material but also to practice thinking like a scientist.

SCIENTIFIC ARGUMENT

Why was it critical in testing the Milankovitch hypothesis to determine the ages of ocean sediment?
Ocean floors accumulate sediment year after year in thin layers. Scientists can drill into the ocean floor and collect cores of those sediment layers, and from chemical tests they can find the temperature of the seawater when each layer was deposited. Those observations of temperature could be used as reality checks for the climate temperatures predicted by the Milankovitch theory, but that meant the ages of the sediment layers had to be determined correctly. When a conflict arose with evidence from Devil's Hole in Nevada, the age determinations of the samples were carefully reexamined and found to be correct.

After reviewing all of the evidence, scientists concluded that the ocean core samples do indeed support the Milankovitch hypothesis. Now construct your own argument. **What might the temperatures recorded at Devil's Hole represent?**

How Do We Know? 3-3

Scientific Arguments

How is a scientific argument different from an advertisement? Advertisements sometimes sound scientific, but they are fundamentally different from scientific arguments. An advertisement is designed to convince you to buy a product. "Our shampoo promises 85 percent shinier hair." The statement may sound like science, but it isn't a complete, honest discussion. "Shinier than what?" you might ask. An advertiser's only goal is a sale.

Scientists construct arguments because they want to test their own ideas and give an accurate explanation of some aspect of nature. For example, in the 1960s, biologist E. O. Wilson presented a scientific argument to show that ants communicate by smells. The argument included a description of his careful observations and the ingenious experiments he had conducted to test his theory. He also considered other evidence and other theories for ant communication. Scientists can include any evidence or theory that supports their claim, but they must observe one fundamental rule of science: They must be totally honest—they must include all of the evidence and all of the theories.

Scientists publish their work in scientific arguments, but they also think in scientific arguments. If, in thinking through his argument, Wilson had found a contradiction, he would have known he was on the wrong track. That is why scientific arguments must be complete and honest. Scientists who ignore inconvenient evidence or brush aside other theories are only fooling themselves.

A good scientific argument gives you all the information you need to decide for yourself whether the argument is correct. Wilson's study of ant communication is now widely understood and is being applied to other fields such as pest control and telecommunications networks.


Eye of Science/Photo Researchers, Inc.

Scientists have discovered that ants communicate with a large vocabulary of smells.

3-3 The Changeable Moon

You have no doubt seen the moon in the sky and noticed that its shape changes from night to night. The cycle of the moon is one of the most obvious phenomena in the sky, and that cycle has been a natural timekeeper since before the dawn of human civilization.

The Motion of the Moon

Just as the planets revolve counterclockwise around the sun as seen from the direction of the celestial north pole, the moon revolves counterclockwise around Earth. Because the moon's orbit is tipped a little over 5° from the plane of Earth's orbit, the moon's path takes it slightly north and then slightly south, but it is always somewhere near the ecliptic.

The moon moves rapidly against the background of the constellations. If you watch the moon for just an hour, you can see it move eastward by slightly more than its angular diameter. In the previous chapter, you learned that the moon is about 0.5° in angular diameter, and it moves eastward a bit more than 0.5° per hour. In 24 hours, it moves 13°. Thus, each night you see the moon about 13° eastward of its location the night before.

As the moon orbits around Earth, its shape changes from night to night in a monthlong cycle.

The Cycle of Phases

The changing shape of the moon as it revolves around Earth is one of the most easily observed phenomena in astronomy. Study **The Phases of the Moon** on pages 34–35 and notice three important points and two new terms:

1. The moon always keeps the same side facing Earth. "The man in the moon" is produced by the familiar features on the moon's near side, but you never see the far side of the moon.
2. The changing shape of the moon as it passes through its cycle of phases is produced by sunlight illuminating different parts of the side of the moon you can see.
3. Notice the difference between the orbital period of the moon around Earth *(sidereal period)* and the length of the lunar phase cycle *(synodic period).* That difference is a good illustration of how your view from Earth is produced by the combined motions of Earth and other heavenly bodies, such as the sun and moon.

You can figure out where the moon will be in the sky by making a moon-phase dial from the middle diagram on page 34. Cover the lower half of the moon's orbit with a sheet of paper and align the edge of the paper to pass through the word "Full"

at the left and the word "New" at the right. Push a pin through the edge of the paper at Earth's North Pole to make a pivot and, under the word "Full," write on the paper "Eastern Horizon." Under the word "New," write "Western Horizon." The paper now represents the horizon you see when you stand facing south. You can set your moon-phase dial for a given time by rotating the diagram behind the horizon-paper. For example, set the dial to sunset by turning the diagram until the human figure labeled "sunset" is standing at the top of the Earth globe; the dial shows that the full moon at sunset would be at the eastern horizon.

The phases of the moon are dramatic, and they have attracted a number of peculiar ideas. You have probably heard a number of **Common Misconceptions** about the moon. Sometimes people are surprised to see the moon in the daytime sky, and they think something has gone wrong! No, the gibbous moon is often visible in the daytime. Although quarter moons and crescent moons can also be located in the daytime sky, they are harder to see when the sun is above the horizon. You may hear people mention "the dark side of the moon," but you will be able to assure them that this is a misconception; there is no permanently dark side. Any location on the moon is sunlit for two weeks and is in darkness for two weeks as the moon rotates in sunlight. Finally, you have probably heard one of the strangest misconceptions about the moon: that people tend to act weird at full moon. (The word *lunatic* comes from *Luna,* the Latin name for the moon.) Actual statistical studies of records from schools, prisons, hospitals, and so on, show that it isn't true. There are always a few people who misbehave; the moon has nothing to do with it.

For billions of years, the man in the moon has looked down on Earth. Ancient civilizations saw the same cycle of phases that you see (■ Figure 3-4), and even the dinosaurs may have noticed the changing phases of the moon. Occasionally, however, the moon displays more complicated moods when it turns copper red in a lunar eclipse.

Lunar Eclipses

A **lunar eclipse** can occur at full moon if the moon moves through the shadow of Earth. Because the moon shines only by reflected sunlight, the moon grows dark while it is crossing through the shadow.

Earth's shadow consists of two parts (■ Figure 3-5). The **umbra** is the region of total shadow. If you were drifting in your spacesuit in the umbra of Earth's shadow, the sun would be completely hidden behind Earth, and you would see no portion of the sun's bright disk. If you drifted into the **penumbra**, however, you would see part of the sun peeking around the edge of Earth, so you would be in partial shadow. In the penumbra, sunlight is dimmed but not extinguished.

Once or twice a year, the orbit of the moon carries it through the umbra of Earth's shadow, and you see a lunar eclipse (■ Figure 3-6). As you watch the eclipse begin, the moon first moves into the penumbra and dims slightly; the deeper it moves into the penumbra, the more it dims. After about an hour, the moon reaches the umbra, and you see the umbral shadow darken part of the moon. It takes about an hour for the moon to enter the umbra completely and become totally eclipsed. **Totality,** the period of total eclipse, may last as long as 1 hour 45 minutes, though the length of totality depends on where the moon crosses the shadow.

When the moon is totally eclipsed, it does not disappear completely. Although it receives no direct sunlight, the moon in the umbra does receive some sunlight that is refracted (bent) through

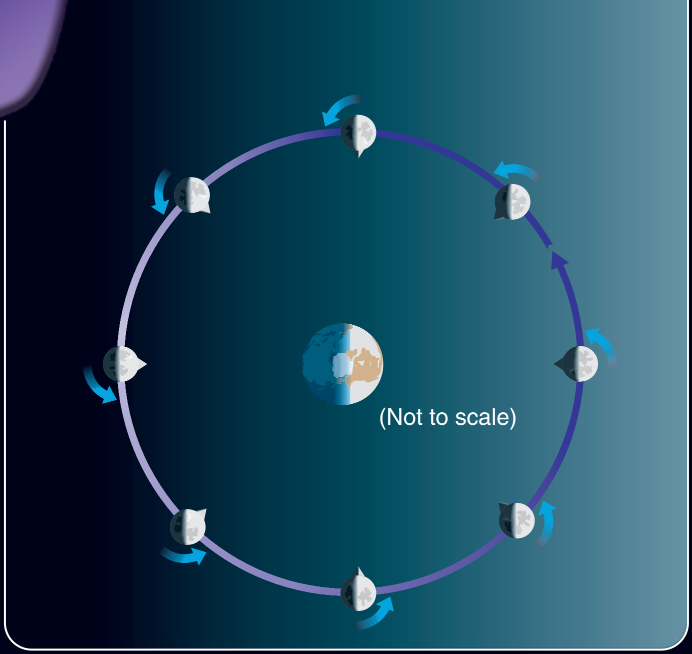
© UC Regents/Lick Observatory

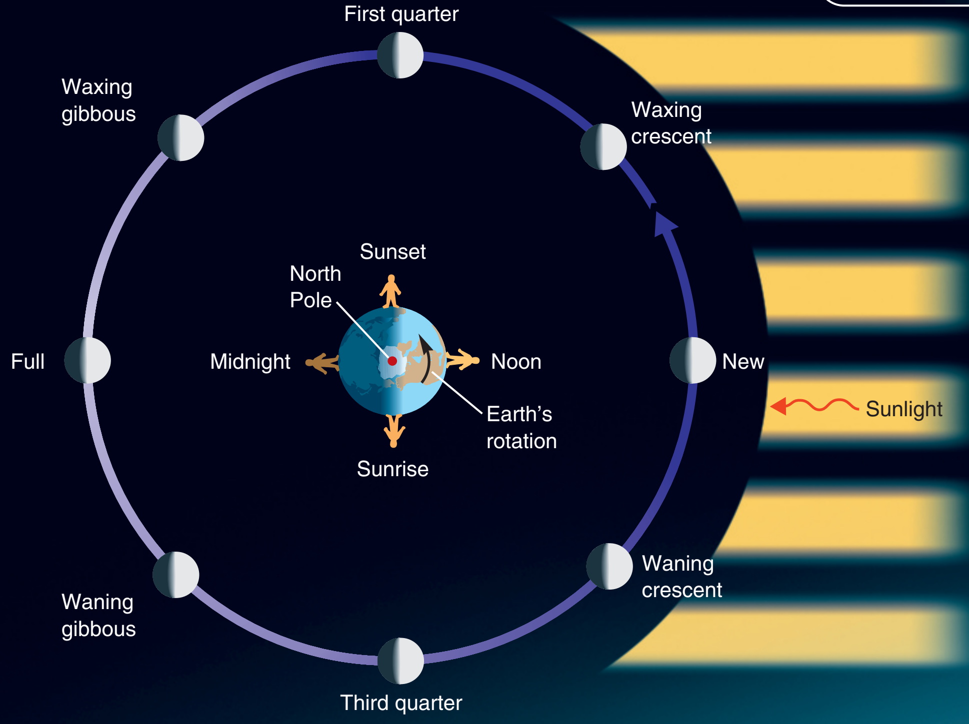
■ **Figure 3-4**

This sequence shows lunar phases night after night through a month from the first crescent at lower right to the last crescent at upper left. Notice that you see the same face of the moon each night, the same mountains, and craters, and plains; it is the changing direction of sunlight that produces the lunar phases.

The Phases of the Moon

1 As the moon orbits Earth, it rotates to keep the same side facing Earth as shown at right. Consequently you always see the same features on the moon, and you never see the far side of the moon. A mountain on the moon that points at Earth will always point at Earth as the moon revolves and rotates.



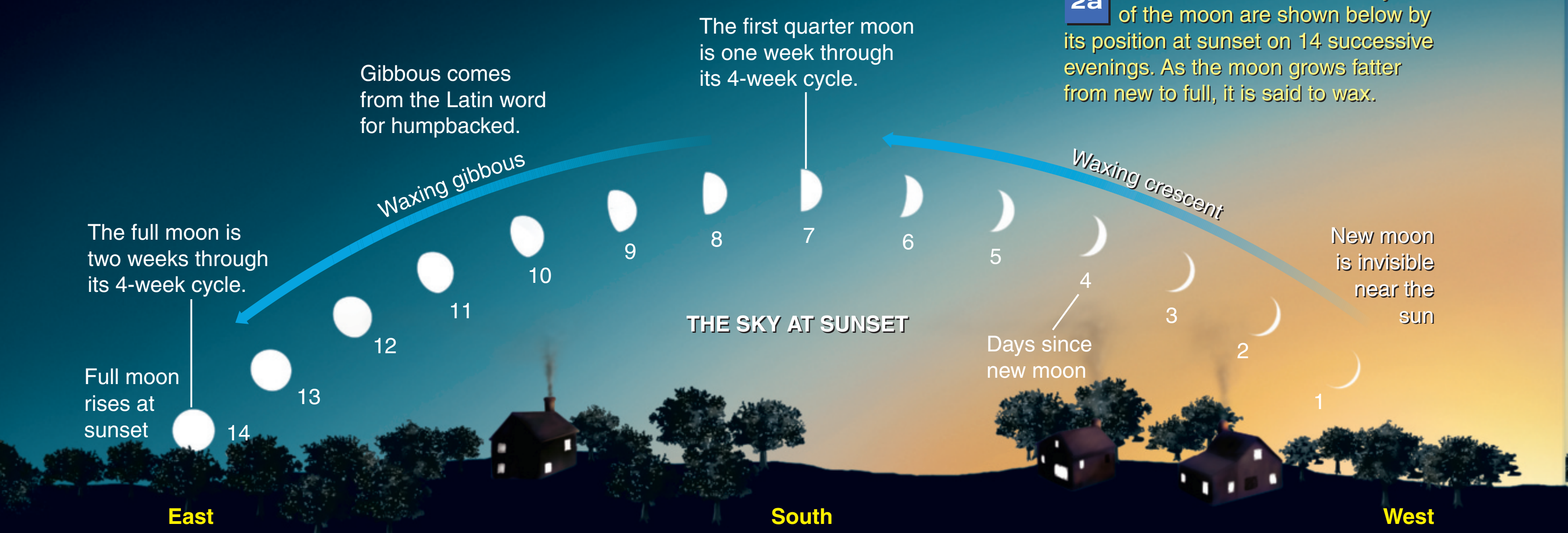


2 As seen at left, sunlight always illuminates half of the moon. Because you see different amounts of this sunlit side, you see the moon cycle through phases. At the phase called "new moon," sunlight illuminates the far side of the moon, and the side you see is in darkness. At new moon you see no moon at all. At full moon, the side you see is fully lit, and the far side is in darkness. How much you see depends on where the moon is in its orbit.

Notice that there is no such thing as the "dark side of the moon." All parts of the moon experience day and night in a month-long cycle.

In the diagram at the left, you see that the new moon is close to the sun in the sky, and the full moon is opposite the sun. The time of day depends on the observer's location on Earth.

2a The first two weeks of the cycle of the moon are shown below by its position at sunset on 14 successive evenings. As the moon grows fatter from new to full, it is said to wax.

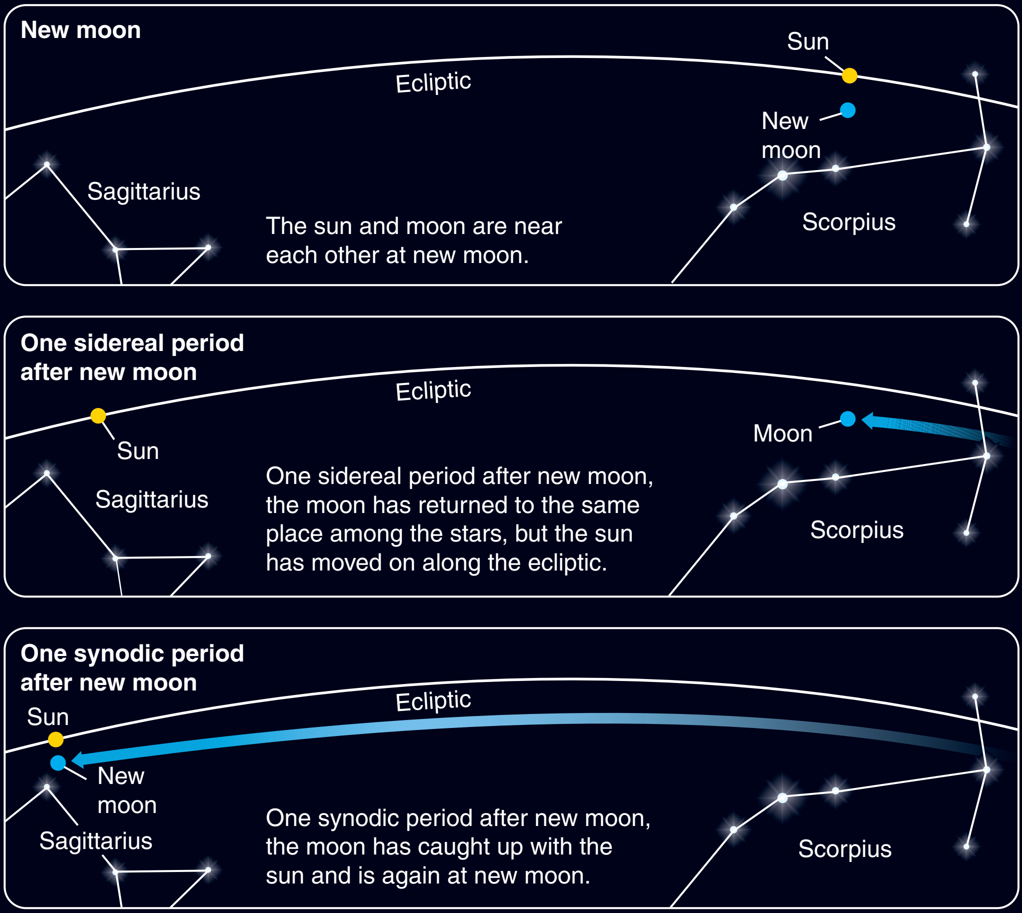


3 The moon orbits eastward around Earth in 27.32 days, its **sidereal period.** This is how long the moon takes to circle the sky once and return to the same position among the stars.

A complete cycle of lunar phases takes 29.53 days, the moon's **synodic period.** (Synodic comes from the Greek words for "together" and "path.")

To see why the synodic period is longer than the sidereal period, study the star charts at the right.

Although you think of the lunar cycle as being about 4 weeks long, it is actually 1.53 days longer than 4 weeks. The calendar divides the year into 30-day periods called months (literally "moonths") in recognition of the 29.53 day synodic cycle of the moon.

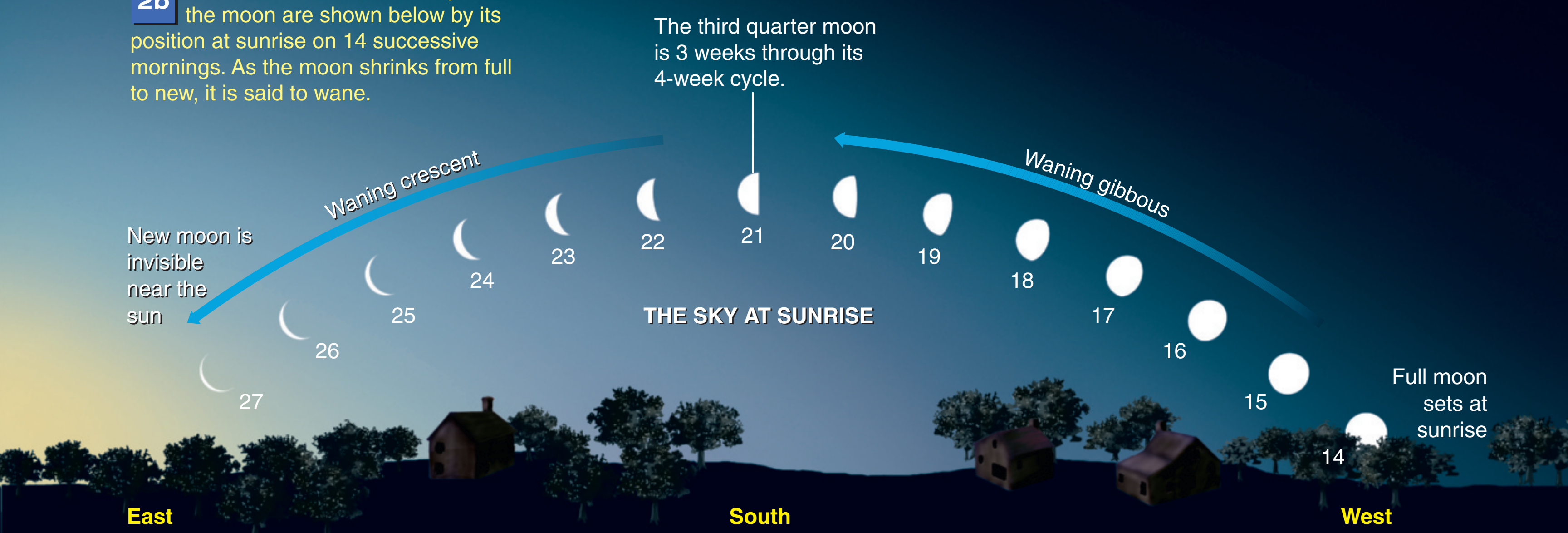


You can use the diagram on the opposite page to determine when the moon rises and sets at different phases.

TIMES OF MOONRISE AND MOONSET

Phase	Moonrise	Moonset
New	Sunrise	Sunset
First quarter	Noon	Midnight
Full	Sunset	Sunrise
Third quarter	Midnight	Noon

2b The last two weeks of the cycle of the moon are shown below by its position at sunrise on 14 successive mornings. As the moon shrinks from full to new, it is said to wane.



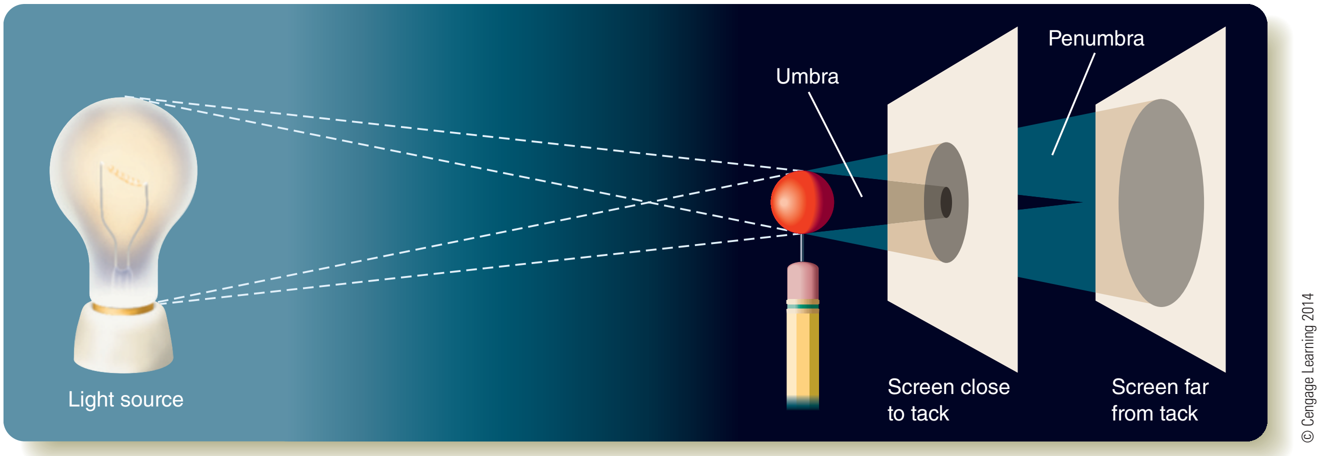


Figure 3-5

The shadows cast by a map tack resemble those of Earth and the moon. The umbra is the region of total shadow; the penumbra is the region of partial shadow.

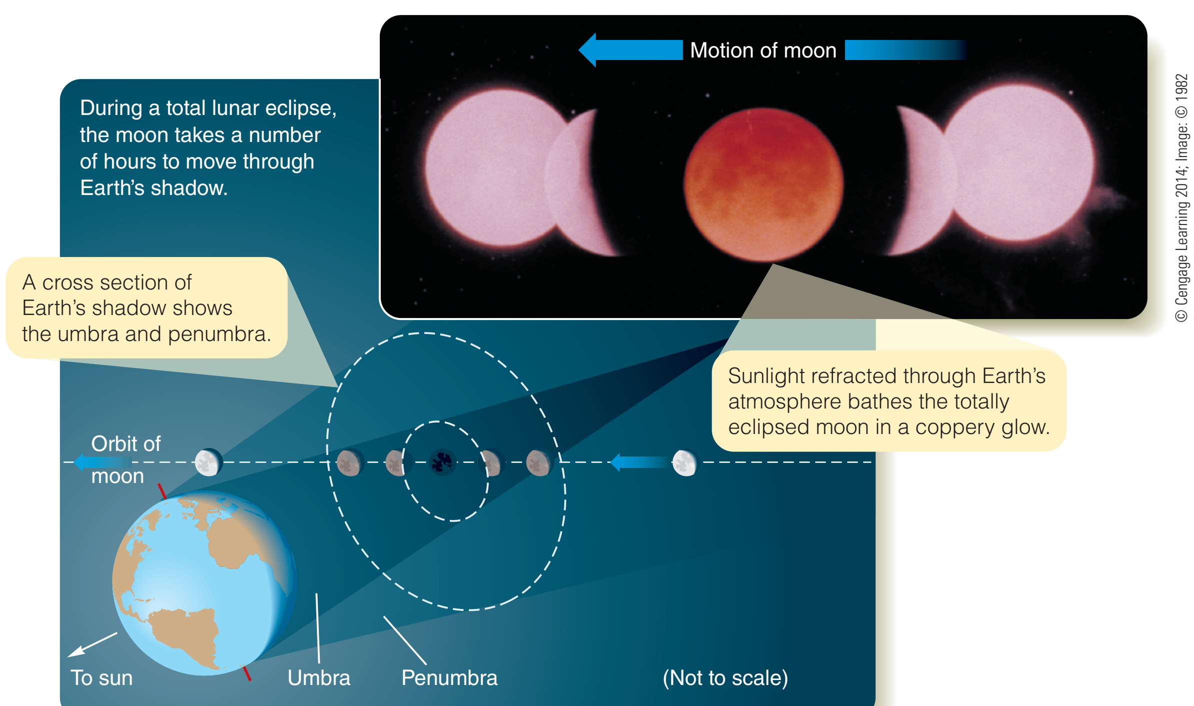


Figure 3-6

During a total lunar eclipse, the moon passes from right to left through Earth's shadow. The multiple-exposure photograph shows the moon at different stages of the eclipse with a longer exposure used to record the moon while it was totally eclipsed. The moon's path appears curved in the photo because of photographic effects.

Earth's atmosphere. If you were on the moon during totality, you would not see any part of the sun because it would be entirely hidden behind Earth. However, you would see Earth's atmosphere illuminated from behind by the sun. The red glow from this ring of "sunsets" and "sunrises" illuminates the moon during totality and makes it glow coppery red, as shown in Figure 3-6.

Lunar eclipses are not always total. If the moon passes a bit too far north or south, it may only partially enter the umbra, and you see a partial lunar eclipse. The part of the moon that remains in the penumbra receives some direct sunlight, and the glare is usually great enough to prevent your seeing the faint coppery glow of the part of the moon in the umbra.

A penumbral lunar eclipse occurs when the moon passes through the penumbra but misses the umbra entirely. Because the penumbra is a region of partial shadow, the moon is only partially dimmed. A penumbral eclipse is not very impressive.

Although there are usually no more than one or two lunar eclipses each year, it is not difficult to see one. You need only be on the dark side of Earth when the moon passes through Earth's shadow. That is, the eclipse must occur between sunset and sunrise at your location. ■ Table 3-1 will allow you to determine which upcoming total and partial lunar eclipses will be visible from your location.

■ Table 3-1 | Total and Partial Eclipses of the Moon, 2013 through 2022*

Date	Time of Mid-Eclipse (GMT)**	Length of Totality (Hr:Min)	Length of Eclipse (Hr:Min)
2013 April 25	20:09	Partial	2:08
2014 April 15	07:47	1:18	3:35
2014 October 8	10:56	0:59	3:20
2015 April 4	12:01	0:05	3:29
2015 September 28	02:48	1:12	3:20
2017 August 7	18:22	Partial	1:55
2018 January 31	13:31	1:16	3:23
2018 July 27	20:23	1:43	3:55
2019 January 21	05:13	1:02	3:17
2019 July 16	21:32	Partial	2:58
2021 May 26	11:20	0:15	3:07
2021 November 19	09:04	Partial	3:28
2022 May 16	04:13	1:25	3:27
2022 November 8	11:00	1:25	3:40

*There are no total or partial lunar eclipses during 2016 or 2020.

**Times are Greenwich Mean Time. Subtract 5 hours for Eastern Standard Time, 6 hours for Central Standard Time, 7 hours for Mountain Standard Time, and 8 hours for Pacific Standard Time. For Daylight Savings Time (mid-March through early November), add 1 hour to Standard Time. Lunar eclipses that occur between sunset and sunrise in your time zone will be visible, and those at midnight will be best placed.

Solar Eclipses

From Earth you can see a phenomenon that is not visible from most planets. It happens because the sun is about 400 times larger than our moon but also, on average, nearly 400 times farther away, so the sun and moon have nearly equal angular diameters of about 0.5°. (See **Reasoning with Numbers 3-1**.) This means that the moon is just the right size to cover the bright disk of the sun and cause a **solar eclipse**. If the moon covers the entire disk of the sun, you see a total eclipse. If it covers only part of the sun, you see a partial eclipse.

At every new moon, the shadow of the moon points toward Earth, but it usually misses. When the moon's shadow does sweep over Earth, the umbra barely reaches Earth and produces a small spot of darkness. The penumbra produces a larger circle of dimmed sunlight (■ Figure 3-8). What you see of the resulting eclipse depends on where you are in those shadows. Standing in that umbral spot, you would be in total shadow, unable to see any part of the sun's bright surface, and the eclipse would be total. If instead you were located outside the umbra, in the penumbra, you would see part of the sun peeking around the edge of the moon, and the eclipse would be partial. Of course, if you are outside the penumbra, you would see no eclipse at all.

Because of the orbital motion of the moon and the rotation of Earth, the moon's shadow sweeps rapidly across Earth in a long, narrow path of totality. If you want to see a total solar eclipse, you must be in the path of totality. When the umbra of the moon's shadow sweeps over you, you see one of the most dramatic sights in the sky—a total eclipse of the sun.

The eclipse begins as the moon slowly crosses in front of the sun. It takes about an hour for the moon to cover the solar disk, but as the last sliver of sun disappears behind the moon, only the glow of the sun's outer atmosphere is visible (■ Figure 3-9) and darkness falls in a few seconds. Automatic streetlights come on, drivers of cars turn on their headlights, and birds go to roost. The sky becomes so dark you can even see the brighter stars.

The darkness lasts only a few minutes because the umbra is never more than 270 km (170 miles) in diameter and sweeps across Earth's surface at over 1600 km/hr (1000 mph). The sun cannot remain totally eclipsed for more than 7.5 minutes, and the average period of totality lasts only 2 or 3 minutes.

The brilliant surface of the sun is called the **photosphere**, and when the moon covers the photosphere, you can see the fainter **chromosphere**, the higher layers of the sun's atmosphere, glowing a bright pink. Above the chromosphere you see the **corona**, the sun's outer atmosphere. The corona is a low-density hot gas that glows with a pale white color. Streamers caused by the solar magnetic field streak the corona, as may be seen in the last frame of Figure 3-9. The chromosphere is often marked by eruptions on the solar surface called **prominences** (■ Figure 3-10a). The corona, chromosphere, and prominences are visible only when the brilliant photosphere is covered. As soon as part of the photosphere

Reasoning with Numbers | 3-1

The Small-Angle Formula

Figure 3-7 shows the relationship between the linear diameter, angular diameter, and distance of an object, in this case the moon. Linear diameter is simply the distance between an object's opposite sides. You are referring to linear diameter when you order a 16-inch pizza—the pizza is 16 inches across. The average linear diameter of the moon is 3470 km (2160 mi). The angular diameter of an object is the angle formed by lines extending toward you from opposite edges of the object, meeting at your eye. Clearly, the farther away a given object is, the smaller its angular diameter.

To find the angular diameter of the moon, you need to use the small-angle formula. It allows you to calculate the angular diameter of any object, whether it is a pizza or the moon. This formula is very important in astronomy, and you will encounter its use many times in later chapters:

$$\frac{\text{angular diameter (in arc seconds)}}{2.06 \times 10^5} = \frac{\text{linear diameter}}{\text{distance}}$$

In the small-angle formula, you must always use the same units for distance and linear diameter. The version of the formula shown here has arc seconds as the unit of angular diameter. You can use this formula to find any one of the three quantities involved (linear diameter, angular diameter, distance) if you know the other two, by cross-multiplying.

Example: The moon's linear diameter is 3470 km, and its average distance from Earth is 384,000 km. What is its angular diameter? **Solution:** Because the moon's linear diameter and distance are both given in the same units, kilometers, you can put them directly into the small-angle formula:

$$\frac{\text{angular diameter}}{2.06 \times 10^5} = \frac{3470 \text{ km}}{384{,}000 \text{ km}}$$

The resulting angular diameter is 1870 arc seconds to three digits' precision, which equals 31 arc minutes—about 0.5°.

* The number 2.06×10^5 in the formula converts the angle units from radians to arc seconds.

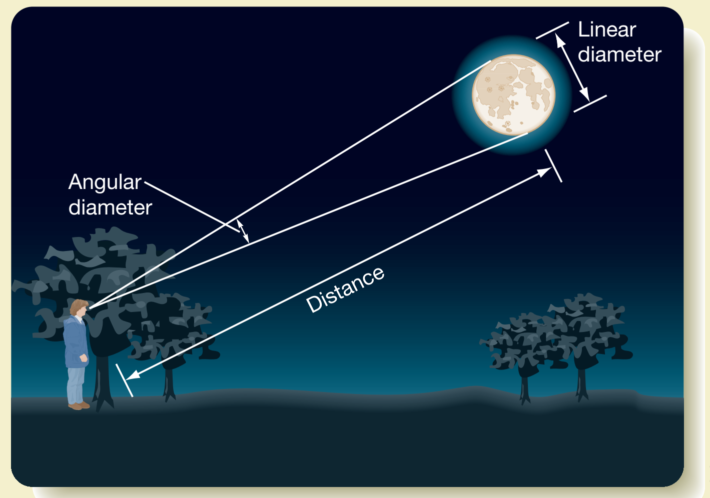


Figure 3-7

The three quantities related by the small-angle formula. Angular diameter is given in arc seconds in the formula. Distance and linear diameter must be expressed in the same units.

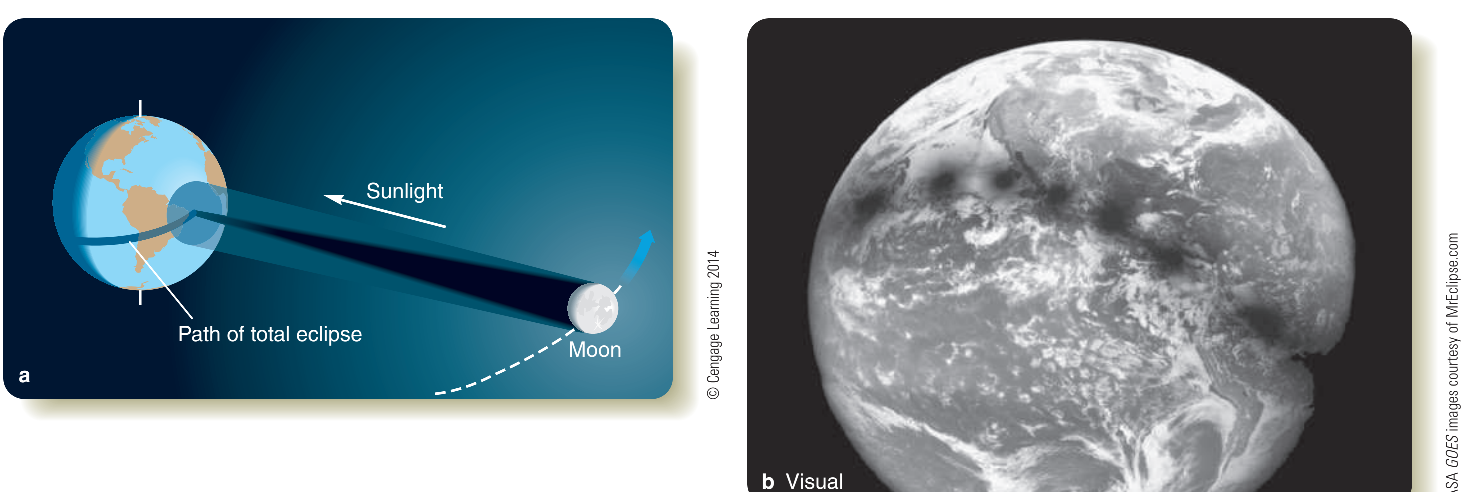
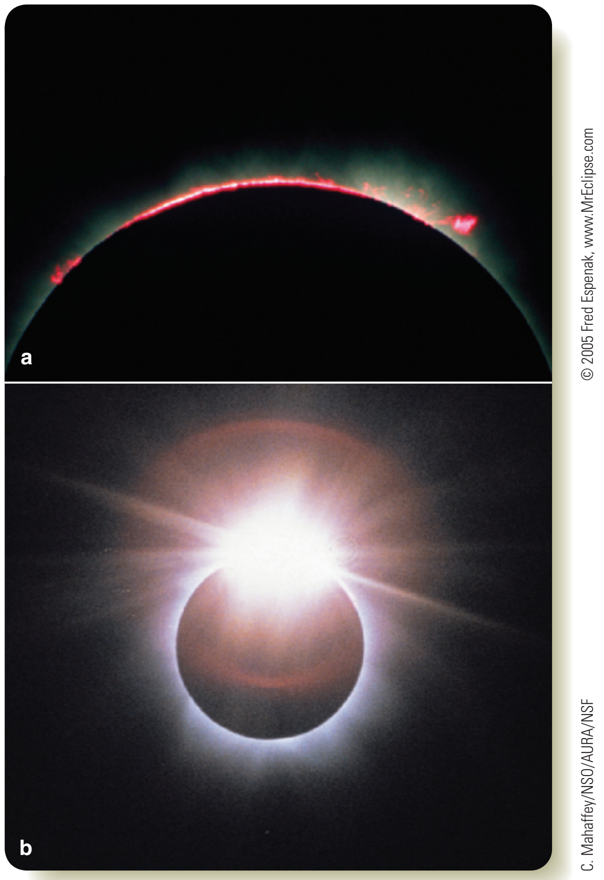


Figure 3-8

(a) The umbra of the moon's shadow sweeps from west to east across Earth, and observers in the path of totality see a total solar eclipse. Those outside the umbra but inside the penumbra see a partial eclipse. (b) Eight photos made by a weather satellite during a total solar eclipse have been combined to show the moon's shadow moving across Mexico, Central America, and Brazil.

■ **Figure 3-9**

This sequence of photos shows the first half of a total solar eclipse.



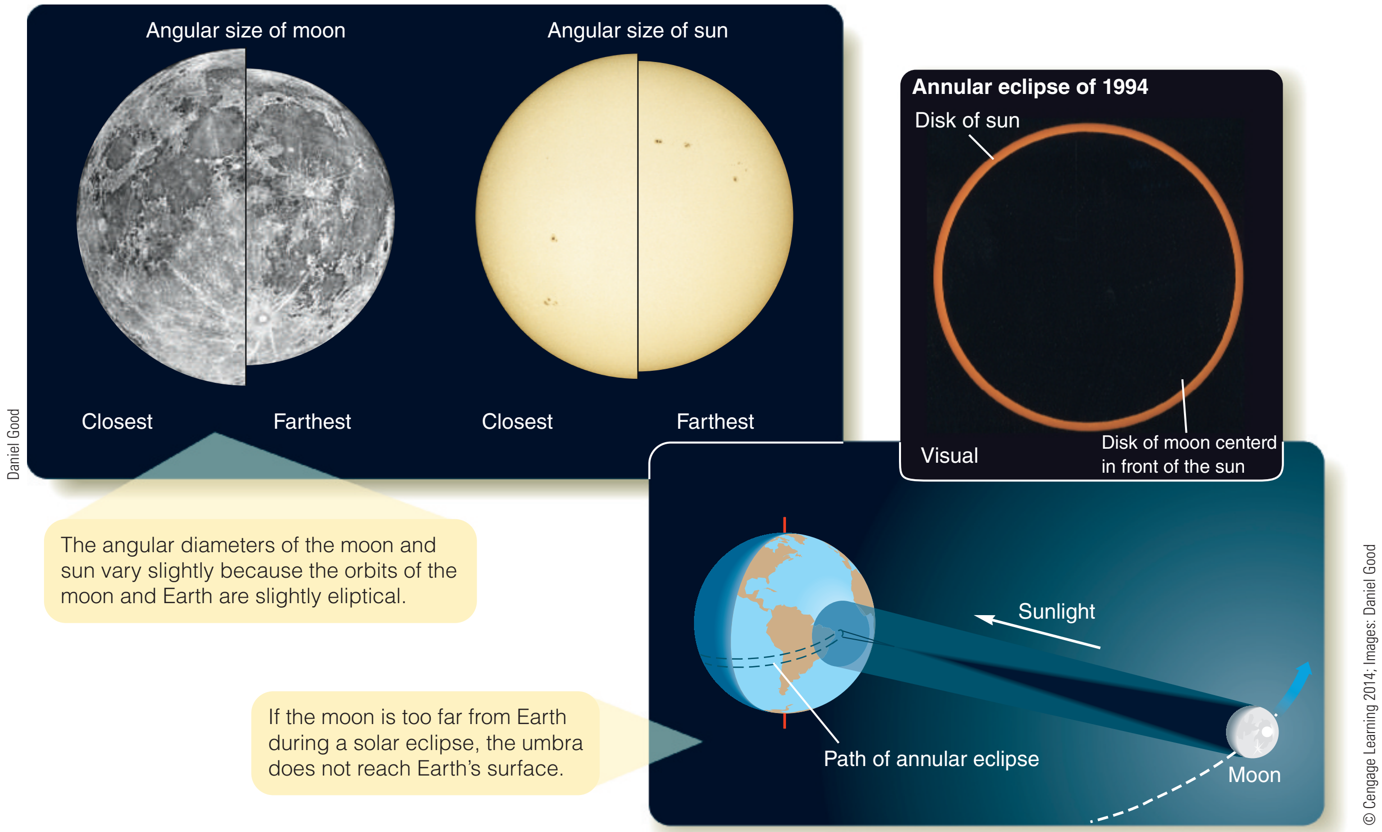
■ **Figure 3-10**

(a) During a total solar eclipse, the moon covers the photosphere, and the ruby red chromosphere and prominences are visible. Only the lower corona is visible in this image. (b) The diamond ring effect can sometimes occur momentarily at the beginning or end of totality if a small segment of the photosphere peeks out through a valley at the edge of the lunar disk.

reappears, the fainter corona, chromosphere, and prominences vanish in the glare, and totality is over. The moon moves on in its orbit, and in an hour the sun is completely visible again.

Just as totality begins or ends, a small part of the photosphere can peek out from behind the moon through a valley at the edge of the lunar disk. Although it is intensely bright, such a tiny bit of the photosphere does not completely drown out the fainter corona, which forms a silvery ring of light, with the brilliant spot of photosphere gleaming like a diamond (Figure 3-10b). This **diamond-ring effect** is one of the most spectacular of astronomical sights, but it is not visible during every solar eclipse. Its occurrence depends on the exact orientation and motion of the moon.

The moon's angular diameter changes depending on where it is around its slightly elliptical orbit. When it is near **perigee**, its point of closest approach to Earth, it looks a little bit larger than when it is near **apogee**, the most distant point in its orbit. Furthermore, Earth's orbit is also slightly elliptical, so the Earth–sun distance varies, and that changes the angular diameter of the solar disk by a few percent (■ Figure 3-11). If the moon is in the farther part of its orbit during totality, its angular diameter will be less than the angular diameter of the sun, and when that happens,



■ **Figure 3-11**

Because the angular diameter of the moon and sun vary slightly, the disk of the moon is sometimes too small to cover the disk of the sun. This means the umbral shadow of the moon does not reach Earth's surface, and the eclipse is annular. From Earth, you see an annular eclipse because the moon's angular diameter is smaller than the angular diameter of the sun. In the photograph of the annular eclipse of 1994, the dark disk of the moon is almost exactly centered on the bright disk of the sun.

you see an **annular eclipse**, a solar eclipse in which a ring (or annulus) of the photosphere is visible around the disk of the moon. Because a portion of the brilliant photosphere remains visible, it never quite gets dark, and you can't see the prominences, chromosphere, and corona (Figure 3-11).

A list of future total and annular eclipses of the sun is given in ■ Table 3-2. If you plan to observe a solar eclipse, remember that the sun is bright enough to burn your eyes and cause permanent damage if you look at it directly. It is a **Common Misconception** that sunlight during an eclipse is somehow extra dangerous. Sunlight is bright enough to burn your eyes any day of the year, whether there is an eclipse or not. Only during totality, while the brilliant photosphere is entirely hidden, is it safe to look directly at the eclipse. See ■ Figure 3-12 for a safe way to observe the partially eclipsed sun.

Predicting Eclipses

Why should you consider eclipse prediction when you can just look them up on the web (http://eclipse.gsfc.nasa.gov/eclipse.html)? One reason is that eclipses occur in a cycle, and as you study eclipse prediction you will see how scientists study cyclic events. A second reason is that the cycle of eclipses is as beautiful as the pattern in a snowflake. The cycle of eclipses will go on around you through your entire life, so you should make it part of your world and enjoy it.

Precise eclipse predictions call for sophisticated calculations, but you can make general predictions by thinking about the geometry of an eclipse and the cyclic motions of the sun and moon.

Solar eclipses occur when the moon passes between Earth and the sun, that is, when the lunar phase is new moon. Lunar eclipses occur at full moon. However, you don't see eclipses at every new moon or full moon. Why not? That's the key question. The answer is that the moon's orbit is tipped a few degrees to the plane of Earth's orbit, so at most new or full moons, the shadows miss, as you can see in the lower part of ■ Figure 3-13. If the shadows miss, there are no eclipses.

For an eclipse to occur, the moon must be passing through the plane of Earth's orbit. The points where it passes through

■ **Table 3-2** | **Total and Annular Eclipses of the Sun, 2013 through 2022***

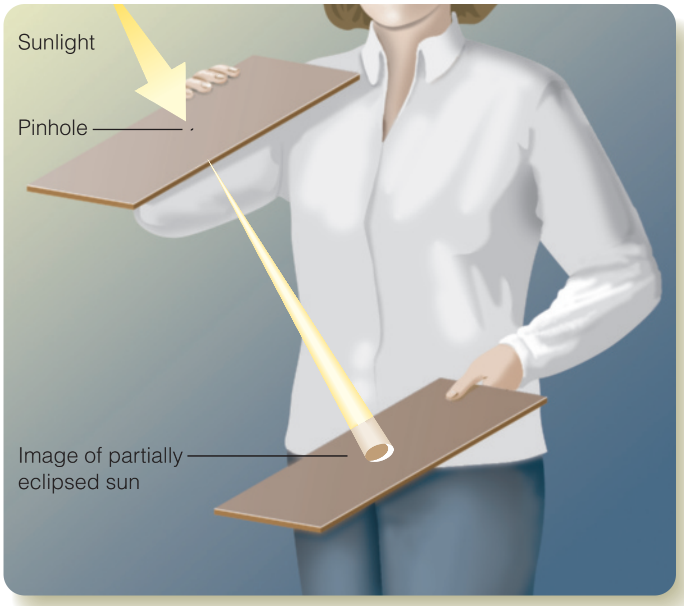
Date	Total/Annular (T/A)	Time of Mid-Eclipse** (GMT)	Maximum Length of Total or Annular Phase (Min:Sec)	Region of Visibility
2013 May 10	A	00:26	6:03	Australia, Pacific
2013 November 3	AT	12:48	1:40	Atlantic, Africa
2015 March 20	T	09:47	2:47	North Atlantic, Arctic
2016 March 9	T	01:58	4:09	Indonesia, Pacific
2016 September 1	A	09:08	3:06	Atlantic, Africa, Indian Ocean
2017 February 26	A	14:55	0:44	South Pacific, South America, Africa
2017 August 21	T	18:27	2:40	Pacific, USA,* Atlantic
2019 July 2	T	19:24	4:33	Pacific, South America
2019 December 26	A	05:19	3:39	Southeast Asia, Pacific
2021 June 10	A	10:43	3:51	North America, Arctic
2021 December 4	T	07:35	1:54	Antarctica, South Atlantic

*The next major total solar eclipse visible from the United States will occur on August 21, 2017, when the path of totality will cross the United States from Oregon to South Carolina.

**There are no total or annular solar eclipses in 2014, 2018, and 2022.

***Times are Greenwich Mean Time. Subtract 5 hours for Eastern Standard Time, 6 hours for Central Standard Time, 7 hours for Mountain Standard Time, and 8 hours for Pacific Standard Time. For Daylight Savings Time (mid-March through early November), add 1 hour to Standard Time.

© Cengage Learning 2014



© Cengage Learning 2014

■ **Figure 3-12**

A safe way to view the partial phases of a solar eclipse. Use a pinhole in a card to project an image of the sun on a second card. The greater the distance between the cards, the larger (and fainter) the image will be.

the plane of Earth's orbit are called the **nodes** of the moon's orbit, and the line connecting these is called the line of nodes. In other words, the planes of the two orbits intersect along the line of nodes. The moon crosses its nodes every month, but eclipses can occur only if the moon is also new or full. That can happen twice a year when the line of nodes points toward the sun, and for a few weeks eclipses are possible at new moons and full moons (Figure 3-13). These intervals when eclipses are possible are called eclipse seasons, and they occur about six months apart.

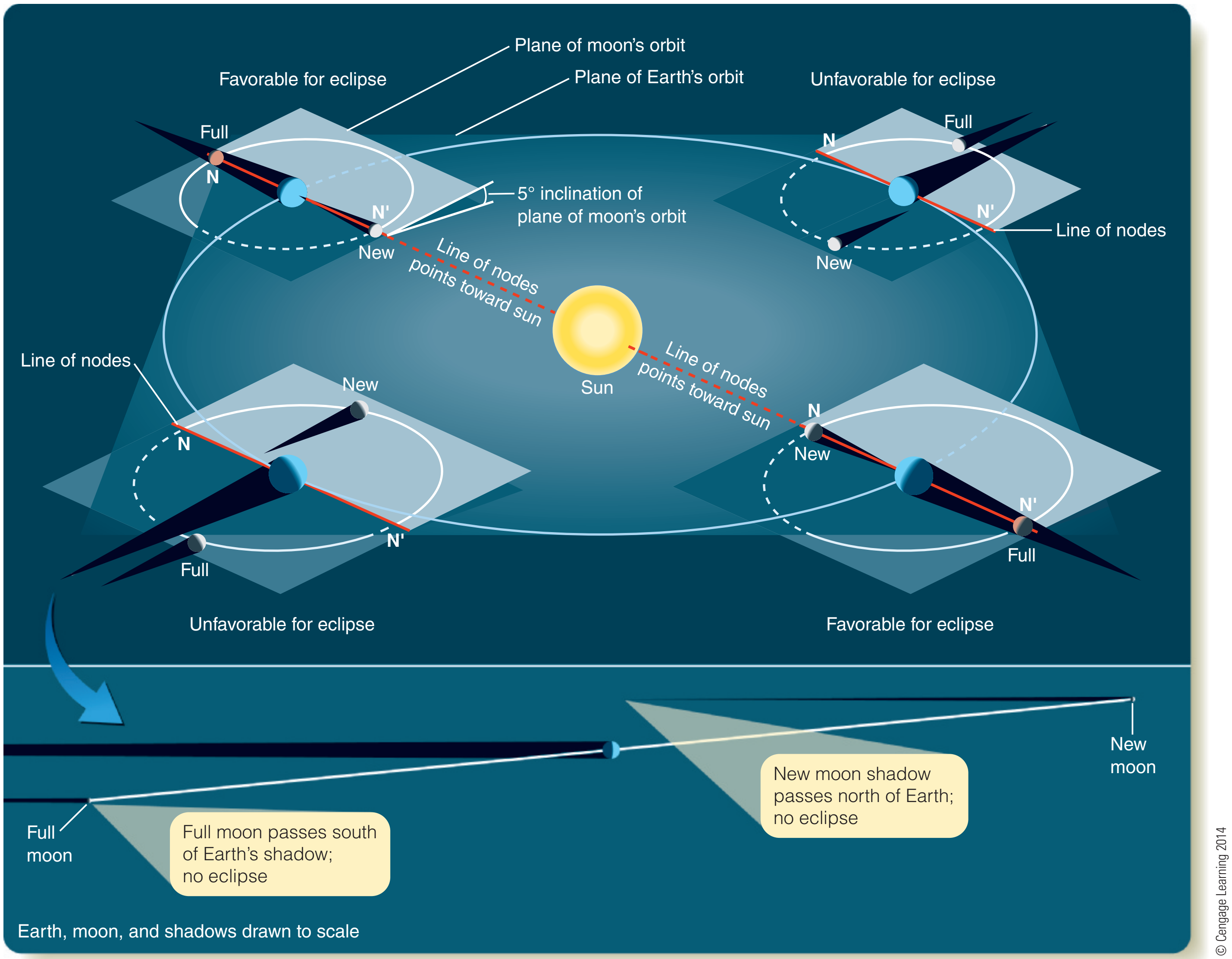


Figure 3-13

The moon's orbit is tipped a bit over 5° to Earth's orbit. The nodes N and N´ are the points where the moon passes through the plane of Earth's orbit. If the line of nodes does not point at the sun, the long narrow shadows miss, and there are no eclipses at new moon and full moon. At those parts of Earth's orbit where the line of nodes points toward the sun, eclipses are possible at new moon and full moon.

If the moon's orbit were fixed in space, the eclipse seasons would always occur at the same times each year. The moon's orbit precesses, however, because of the gravitational pull of the sun on the moon, and the precession slowly changes the direction of the line of nodes. The line turns gradually westward, making one complete rotation in 18.61 years. As a result, the eclipse seasons occur about three weeks earlier each year. Many ancient peoples noticed this pattern and could guess which full and new moons were likely to produce eclipses.

Another way the ancients predicted eclipses was to notice that the pattern of eclipses repeats every 6585.3 days—the **Saros cycle**. After one Saros, the sun, moon, and nodes have circled the sky many times and finally returned to the same arrangement they occupied when the Saros began. Then the cycle of eclipses begins to repeat. One Saros equals 18 years 11⅓ days. Because of the extra third of a day, an eclipse visible in North America will recur after one Saros, but it will be visible one third of the way around the world in the

North Pacific. Once ancient astronomers recognized the Saros cycle, they could predict eclipses from records of previous eclipses.

SCIENTIFIC ARGUMENT

What would astronauts on the moon observe while people on Earth were seeing a total lunar eclipse?

This scientific argument requires that you change your point of view and imagine seeing an event from a new location. Remember that when you see a total lunar eclipse, the full moon is passing through Earth's shadow. Astronauts standing on the moon would look up and see Earth crossing in front of the brilliant sun. Sunlight would begin to dim as the moon entered Earth's penumbra. The visible part of the sun would grow narrower and narrower until it vanished entirely behind Earth, and the astronauts would be left standing in the dark as the moon carried them through the umbra of Earth's shadow. Except for faint starlight, their only light would come from the glow of Earth's atmosphere lit from behind, a red ring around the dark disk of Earth made up of every sunset and sunrise. The red light from Earth's atmosphere would bathe the dusty plains and mountains of the moon in a copper red glow. The astronauts would have a cold and tedious wait for the sun to reemerge from behind Earth, but they would see a lunar eclipse from a new and dramatic vantage point.

Imagining the same event from different points of view can help you sort out complex geometries. Now change your argument slightly and imagine the eclipse once again. **If Earth had no atmosphere, how would this eclipse look different as viewed from Earth and from the moon?**

What Are We? Scorekeepers

The rotation and revolution of Earth produce the cycles of day and night and winter and summer, and we have evolved to live within those cycles. One theory holds that we sleep at night because dozing in the back of a cave (or in a comfortable bed) is safer than wandering around in the dark. The night is filled with predators, so sleeping may keep us safe. Our bodies depend on that cycle of light and dark. People who live and work in the Arctic or Antarctic where the cycle of day and night does not occur can suffer psychological problems from the lack of the daily cycle.

The cycle of the seasons controls the migration of game and the growth of crops, so cultures throughout history have followed the motions of the sun along the ecliptic with special reverence. The people who built Stonehenge were marking the summer solstice sunrise because it was a moment of power, order, and promise in the cycle of their lives.

The moon's cycles mark the passing days and divide our lives into weeks and months. In a Native American story, Coyote gambles with the sun to see if the sun will continue to warm Earth, and the moon keeps score. The moon is a symbol of regularity, reliability, and dependability. It is the trustworthy scorekeeper counting out your days and months.

Like the ticking of a cosmic clock, the passing weeks, months, and seasons mark the passage of time on Earth, but, as you have seen, the cycle of the seasons is also affected by longer period changes in the motion of Earth. Ice ages come and go, and Earth's climate follows complex cycles in ways we do not entirely understand. If you don't feel quite as secure as you did when you started this chapter, then you are catching on. Astronomy tells us that Earth is a beautiful world, but it is also a complicated, spinning planet. Our clocks, calendars, and lives count the passing cycles in the sky.

Study and Review

Summary

- The **rotation (p. 24)** of Earth on its axis produces the cycle of day and night, and the **revolution (p. 24)** of Earth around the sun produces the cycle of the year.
- Because Earth orbits the sun, the sun appears to move eastward along the **ecliptic (p. 25)** through the constellations, completing a circuit of the sky in a year. Because the ecliptic is tipped 23.4° to the celestial equator, the sun spends half the year in the northern celestial hemisphere and half in the southern celestial hemisphere.
- In the summer, the sun is above the horizon longer and shines more directly down on the ground, producing warmer weather. In the winter, the sun is above the horizon for a shorter time and shines on the ground at an angle. That produces colder weather. Summer occurs in the United States and Europe in July, August, and September, while the sun is in the northern sky, but the seasons are reversed in Earth's southern half relative to the Northern Hemisphere. Australia, for example, experiences summer in January, February, and March when the sun is in the southern sky.
- The beginnings of the seasons are marked by the **vernal equinox (p. 26)**, the **summer solstice (p. 26)**, the **autumnal equinox (p. 26)**, and the **winter solstice (p. 26)**.
- Earth is slightly closer to the sun at **perihelion (p. 27)** in January and slightly farther away from the sun at **aphelion (p. 27)** in July. This has almost no effect on the seasons. These names refer to the seasons in Earth's Northern Hemisphere.
- The planets move generally eastward along the ecliptic, and all but Uranus and Neptune are visible to the unaided eye, looking like stars. Mercury and Venus never wander far from the sun and are sometimes visible in the evening sky after sunset or in the dawn sky before sunrise.
- Planets visible in the sky at sunset are traditionally called **evening stars (p. 25)**, and planets visible in the dawn sky are called **morning stars (p. 25)**, even though they are not really stars.
- The locations of the sun and planets along the **zodiac (p. 28)** are diagramed in a **horoscope (p. 28)**, which is the bases for the ancient **pseudoscience (p. 29)** known as astrology.
- According to the **Milankovitch hypothesis (p. 28)**, slow changes in the shape of Earth's orbit, axis tilt, and axis orientation can alter the planet's heat balance and cause the cycles of glacial advances and retreats. Evidence found in seafloor sediments and other locations support the hypothesis, which is widely accepted today.
- Scientists routinely test their own ideas by organizing theory and evidence into a **scientific argument (p. 31)**.
- The moon orbits eastward around Earth once a month and rotates on its axis, keeping the same side facing Earth throughout the month.
- Because you see the moon by reflected sunlight, its shape appears to change as it orbits Earth and sunlight illuminates different amounts of the side you can see.
- The lunar phases wax from new moon to first quarter to full moon and wane from full moon to third quarter to new moon.
- A complete cycle of lunar phases takes 29.53 days, which is known as the moon's **synodic period (p. 35)**. The **sidereal period (p. 35)** of the moon—its orbital period with respect to the stars—is 27.32 days, more than 2 days shorter.
- If a full moon passes through Earth's shadow, sunlight is cut off, and the moon darkens in a **lunar eclipse (p. 33)**. If the moon fully enters the dark **umbra (p. 33)** of Earth's shadow, the eclipse is total, but if it only grazes the umbra, the eclipse is partial. If the moon enters the partial shadow of the **penumbra (p. 33)** but not the umbra, the eclipse is penumbral.
- During **totality (p. 33)**, the eclipsed moon looks copper red because of sunlight refracted through Earth's atmosphere.
- A **solar eclipse (p. 37)** occurs if a new moon passes between the sun and Earth and the moon's shadow sweeps over Earth's surface along the path of totality. Observers inside the path of totality see a total eclipse, and those just outside the path of totality see a partial eclipse as the penumbra sweeps over their location.
- During a total eclipse, the bright **photosphere (p. 37)** of the sun is covered, and the fainter **corona (p. 37)**, **chromosphere (p. 37)**, and **prominences (p. 37)** become visible.
- Sometimes just as totality begins or ends, the bright photosphere peeks out through a valley at the edge of the lunar disk and produces the **diamond-ring effect (p. 38)**.
- When the moon is near **perigee (p. 38)**, the closest point in its orbit, its angular diameter is large enough to cover the sun's photosphere and produce a total eclipse. But if the moon is near **apogee (p. 39)**, the farthest point in its orbit, it looks too small and can't entirely cover the photosphere. A solar eclipse occurring then would be an **annular eclipse (p. 39)**.
- Looking at the sun is dangerous and can burn the retinas of your eyes. The safest way to observe the partial phases of a solar eclipse is by pinhole projection. Only during totality, when the photosphere is completely hidden, is it safe to look at the sun directly.
- Because the moon's orbit is tipped a few degrees from the plane of Earth's orbit, most full moons pass north or south of Earth's shadow, and no lunar eclipse occurs. Also, most new moons cross north or south of the sun, the moon's shadow does not sweep over Earth, and there is no solar eclipse.
- Eclipses can only occur when a full moon or a new moon occurs near one of the two **nodes (p. 40)** of its orbit, where it crosses the ecliptic. These two eclipse seasons occur about 6 months apart, but move slightly earlier each year. By keeping track of the location of the nodes of the moon's orbit, you could predict which full and new moons were most likely to be eclipsed.
- Eclipses follow a pattern lasting 18 years 11 ⅓ days called the **Saros cycle (p. 42)**. If ancient astronomers understood that pattern, they could predict eclipses.

Study and Review

Review Questions

1. What is the difference between the daily and annual motions of the sun?
2. If Earth did not rotate, could you still define the ecliptic? Why or why not?
3. What would the seasons be like if Earth were tipped 35° instead of 23.4°? What would they be like if Earth's axis were perpendicular to its orbit?
4. Why are the seasons reversed in the Southern Hemisphere relative to the Northern Hemisphere?
5. How could small changes in the inclination of Earth's axis affect world climate?
6. Do the phases of the moon look the same from every place on Earth, or is the moon full at different times as seen from different locations?
7. What phase would Earth be in if you were on the moon when the moon was full? At first quarter? At waning crescent?
8. Why have most people seen a total lunar eclipse, while few have seen a total solar eclipse?
9. Why isn't there an eclipse at every new moon and at every full moon?
10. Why is the moon red during a total lunar eclipse?
11. Why should the eccentricity of Earth's orbit make winter in the Northern Hemisphere different from winter in the Southern Hemisphere?
12. **How Do We Know?** What are the main characteristics of a pseudoscience? Can you suggest other examples?
13. **How Do We Know?** Why would it be appropriate to refer to evidence as the reality checks in science?
14. **How Do We Know?** Why must a scientific argument dealing with some aspect of nature include all of the evidence?

Discussion Questions

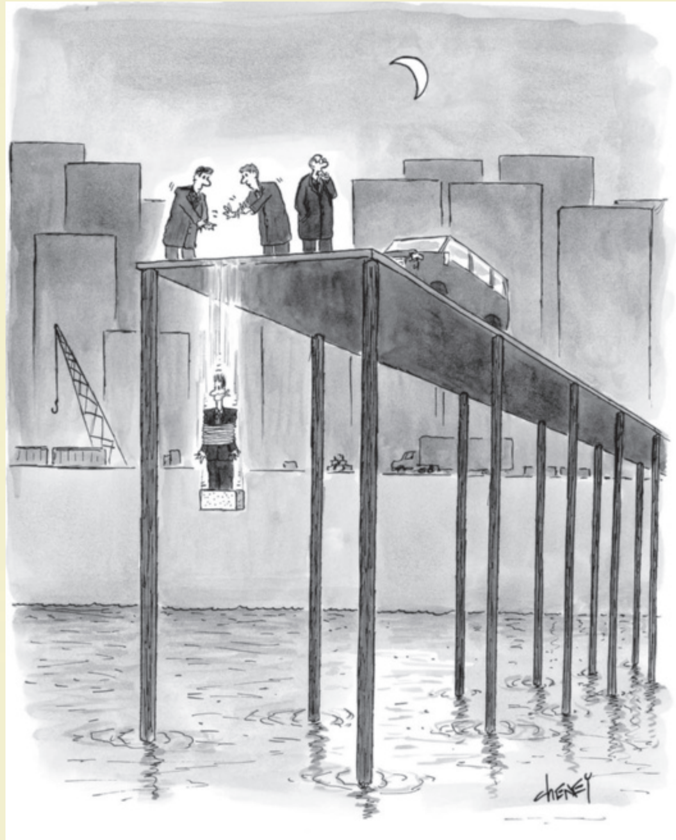
1. Do planets orbiting other stars have ecliptics? Could they have seasons?
2. Why would it be difficult to see prominences if you were on the moon during a total lunar eclipse?
3. Each year the eclipse seasons occur about 19 days earlier. Can you detect that in Tables 3-1 and 3-2?

Problems

1. Given that Earth is about 4.6 billion (4.6×10^9) years old, how many precessional cycles have occurred?
2. Identify the phases of the moon if on March 20 the moon were located at the position the sun is located on: (a) March 20, (b) September 22, (c) June 22, and (d) December 22.
3. Identify the phases of the moon if at sunset in the northern hemisphere the moon were (a) near the eastern horizon, (b) high in the south, (c) in the southeast, (d) in the southwest.
4. About how many days must elapse between first-quarter moon and third-quarter moon?
5. Draw a diagram showing Earth, the moon, and shadows during (a) a total solar eclipse, (b) a total lunar eclipse, (c) a partial lunar eclipse, (d) an annular eclipse.
6. Phobos, one of the moons of Mars, is about 25 km in diameter and orbits about 6000 km above the surface of the planet. What is the angular diameter of Phobos as seen from Mars? (*Hint:* See Reasoning with Numbers 3-1.)
7. A total eclipse of the sun was visible from Canada on July 10, 1972. When did the next eclipse with the same sun-moon-Earth geometry occur? From what part of Earth was it total?
8. When will the eclipse described in Problem 7 next be total as seen from Canada?

Learning to Look

1. The cartoon below shows a crescent moon. Explain why the moon could never look this way.



"Fixing a leak—and you?"

2. The photo below shows the annular eclipse of May 30, 1984. How is it different from the annular eclipse shown in Figure 3-11? Why do you suppose it is different?

Laurence Marschall

Great Debates

1. *Renaming Names in Astronomy.* Typically in astronomy, once a celestial object is named, that name sticks. For example, evening and morning stars are actually planets, not stars; shooting stars are meteors, not stars; planetary nebulae are not planets; pulsars do not pulsate. Is the scientific community doing a disservice to the public by not changing these names to reflect their true meanings? Why or why not?
   a. Choose at least two celestial objects such as those in the list above. Include definitions of the objects (i) listed in the text and (ii) described in your own words. Italicize *your* definitions.
   b. What's the evidence? Find example pictures of your objects and explain the pictures using *your* definitions.
   c. Cite your sources, including page and paragraph numbers in the text.

2. *Colonizing the Moon.* Certain cultures view the moon as a sacred being. Today space travel is commercialized, and in the next several years anyone will be able to travel to space and possibly to the moon. Should commercial spacecrafts and, thus, humans land on the moon? Are humans being insensitive to these cultures that view the moon as a sacred being? If anything, what should be done, and by whom, to recognize these beliefs? Find pictures of other cultures and/or objects that humans may have been insensitive to in a similar way in the past.
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Embed copies of your pictures and explain the pictures.
   c. Cite your sources.

3. *Zodiac.* Thirteen zodiac constellations coincide with the ecliptic, but only twelve of these are associated with a horoscope. Should the horoscope be modified to include the thirteenth zodiacal constellation, Ophiuchus, which is between Scorpius and Sagittarius?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

4. *Mining the Moon.* A candidate states during an election debate that if he is elected president of the United States, he will build a lunar colony to mine minerals from the moon. Should the moon be available as a mining site for minerals? Is a lunar colony a wise use of taxpayer dollars?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

5. *Is Milankovitch Right?* You read in this chapter about the Milankovitch hypothesis. Milutin Milankovitch thought that climate change and ice ages could be caused by the combination of small changes in Earth's orbit, precession, and inclination angle. Did Milankovitch get it wrong? Don't confuse the Milankovitch cycles with global warming. They are quite different.
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

Enhanced Web Assign

ENHANCED WebAssign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Lunar Eclipse
- Seasons
- Path of the Sun
- Shadow
- Small-Angle Formula
- Solar Eclipse
- Solar and Lunar Eclipses

4

The Origin of Modern Astronomy

Guidepost

The preceding chapters gave you a modern view of Earth. You can now imagine how Earth, the moon, and the sun move through space and how that produces the sights you see in the sky. But how did humanity first realize that we live on a planet moving through space? That required the revolutionary overthrow of an ancient and honored theory of Earth's place.

By the 16th century, many astronomers were uncomfortable with the theory that Earth sat at the center of a spherical universe. In this chapter, you will discover how an astronomer named Copernicus changed the old theory, how Galileo Galilei changed the rules of debate, and how Isaac Newton changed humanity's concept of nature. Here you will find answers to four essential questions:

- **How did classical philosophers describe Earth's place in the universe?**
- **How did Copernicus revise that ancient theory?**
- **Why was Galileo condemned by the Inquisition?**
- **How did Isaac Newton change humanity's view of nature?**

This chapter is not just about the history of astronomy. As the astronomers of the Renaissance struggled to understand Earth and the heavens, they invented a new way of understanding nature—a way of thinking that is now called science. Every chapter that follows will use the methods that were invented when Copernicus tried to repair that ancient theory that Earth was the center of the universe.

Galileo's telescope revealed such things as craters on the moon, and he explained how that evidence could be used to test the prevailing Earth-centered model of the universe. He was condemned by the Inquisition in 1633.

Huens, Jean-Leon (1921-82)/National Geographic Image Collection/The Bridgeman Art Library International

How you would burst out laughing, my dear Kepler, if you would hear what the greatest philosopher of the Gymnasium told the Grand Duke about me . . .

FROM A LETTER BY GALILEO GALILEI

NEXT TIME YOU look at the sky, imagine how prehistoric families felt as they huddled around the safety of their fires and looked up at the stars. Astronomy had its beginnings in simple human curiosity about the lights in the sky. As early civilizations developed, great philosophers struggled to understand the movements of the sun, moon, and planets. Later, mathematical astronomers made precise measurements and computed detailed models in their attempts to describe celestial motions. It took hard work and years of effort, but the passions of astronomy gripped some of the greatest minds in history and drove them to try to understand the sky. As you study the history of astronomy, notice that two themes twist through the story.

One theme is the struggle to understand the place of Earth in the universe. It seemed obvious to the ancients that Earth was the center of everything, but today you know that's not true. The debate over the place of Earth involved deep theological questions and eventually led Galileo before the Inquisition.

The second theme is the long and difficult quest to understand planetary motion. Astronomers built more and more elaborate mathematical models, but they still could not predict precisely the motion of the visible planets along the ecliptic. That mystery was finally solved when Isaac Newton described gravity and orbital motion in the late 1600s.

Only a few centuries ago, as astronomers were struggling to understand the sky, they invented a new way of understanding nature—a new way of knowing about the physical world based on the comparison of theories and evidence. Today, that new way of knowing is called science.

4-1 Classical Astronomy

THE GREAT PHILOSOPHERS of ancient Greece wrote about many different subjects, including what they saw in the sky. Those writings became the foundation on which later astronomers built modern astronomy.

The Aristotelian Universe

You have probably heard of the two greatest philosophers of ancient Greece—Plato and Aristotle. Their writings shaped the history of astronomy. Plato (427?–347 BCE) wrote about moral responsibility, ethics, the nature of reality, and the ideals of civil government. His student Aristotle (384–322 BCE) wrote on almost every area of knowledge and is probably the most famous philosopher in history. These two philosophers established the first widely accepted ideas about the structure of the universe.

Science and its methods of investigation did not exist in ancient Greece, so when Plato and Aristotle turned their minds to the problem of the structure of the universe, they made use of a process common to their times—reasoning from first principles. A first principle is something that is held to be obviously true. Once a principle is recognized as true, whatever can be logically derived from it must also be true.

But what was obviously true to the ancients is not so obvious to us today. Study **The Ancient Universe** on pages 48–49 and notice three important ideas and seven new terms that show how first principles influenced early descriptions of the universe and its motions:

1. Ancient philosophers and astronomers accepted as first principles that Earth was located at the center of a *geocentric universe* and that everything in the heavens moved in *uniform circular motion.* They thought it was obvious that Earth did not move because they did not see the shifting of the stars called *parallax.*
2. Notice how the observed motion of the planets did not fit the theory very well. The *retrograde motion* of the planets was very difficult to explain using geocentrism and uniform circular motion.
3. Claudius Ptolemy attempted to explain the motion of the planets mathematically by devising a small circle, the *epicycle,* rotating along the edge of a larger circle, the *deferent,* that enclosed Earth. He even allowed the speed of the planets to vary slightly as they circled a slightly off-center point called the *equant.* In these ways he weakened the principles of geocentrism and uniform circular motion.

Ptolemy lived roughly five centuries after Aristotle in the Greek colony in Egypt, and although Ptolemy accepted the Aristotelian universe, he was interested in a different problem—the motion of the planets. He was a brilliant mathematician, and he used his talents to create a mathematical description of the motions he saw in the heavens. For him, first principles took second place to mathematical precision.

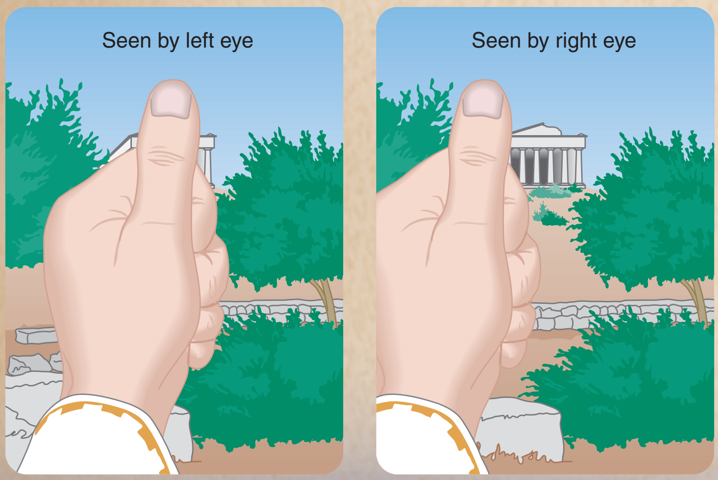
Aristotle's universe, as embodied in Ptolemy's mathematical model, dominated ancient astronomy, but it was wrong. The universe is not geocentric, and the planets don't follow circles at uniform speeds. At first the Ptolemaic system predicted the positions of the planets well; but, as centuries passed, errors accumulated. Astronomers tried to update the system, computing new constants and adjusting epicycles. In the middle of the 13th century, a team of astronomers supported by King Alfonso X of Castile studied the *Almagest* for 10 years. Although they did not revise the theory very much, they simplified the calculation of the positions of the planets using the Ptolemaic system and published the result as

1 For 2000 years, the minds of astronomers were shackled by a pair of ideas. The Greek philosopher Plato argued that the heavens were perfect. Because the only perfect geometrical shape is a sphere, which carries a point on its surface around in a circle, and because the only perfect motion is uniform motion, Plato concluded that all motion in the heavens must be made up of combinations of circles turning at uniform rates. This idea was called **uniform circular motion.**

Plato's student Aristotle argued that Earth was imperfect and lay at the center of the universe. Such a model is known as a **geocentric universe.** His model contained 55 spheres turning at different rates and at different angles to carry the seven known planets (the moon, Mercury, Venus, the sun, Mars, Jupiter, and Saturn) across the sky.

Aristotle was known as the greatest philosopher in the ancient world, and for 2000 years his authority chained the minds of astronomers with uniform circular motion and geocentrism. See the model at right.

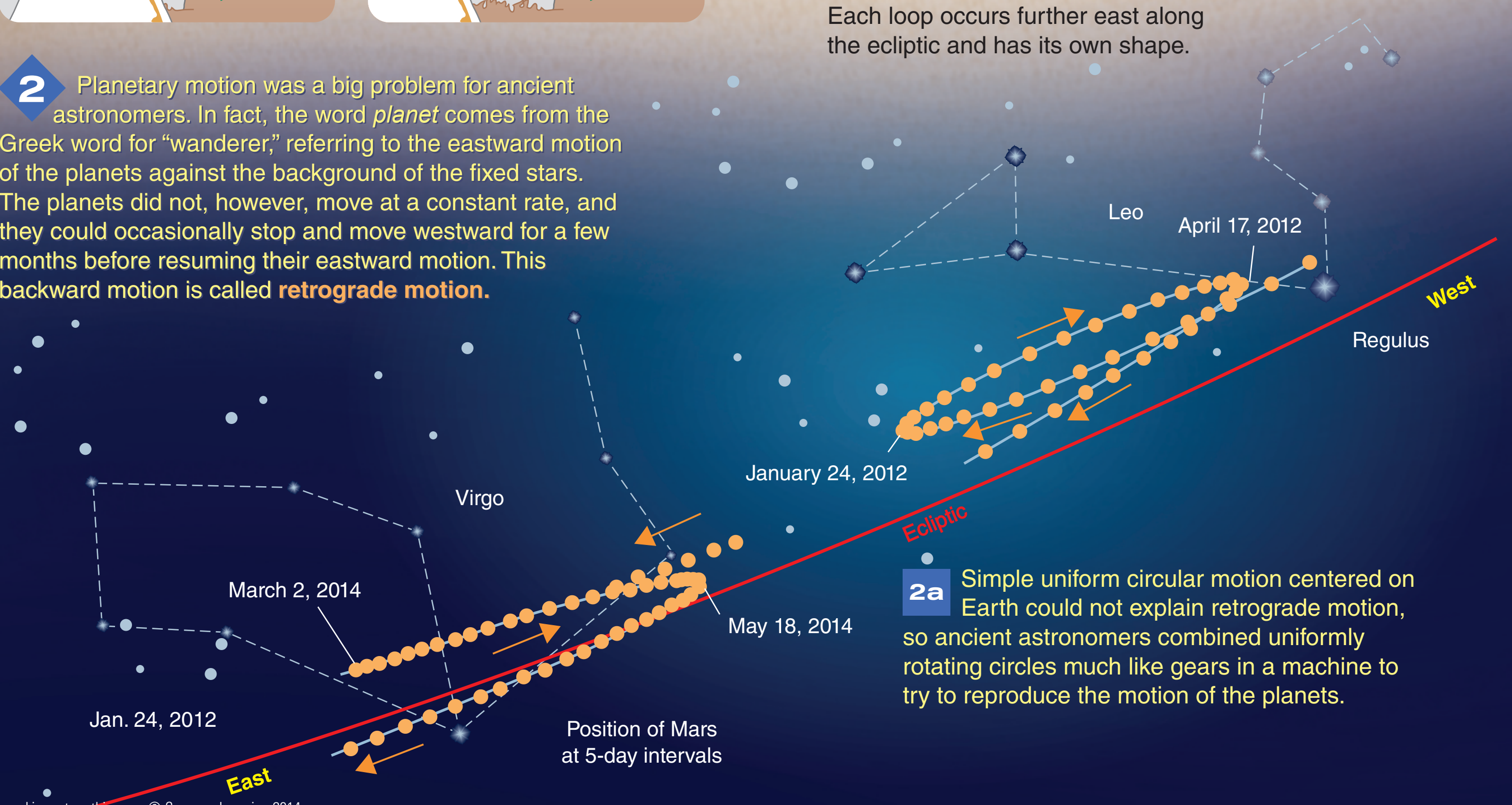
From *Cosmographica* by Peter Apian (1539).



1a Ancient astronomers believed that Earth did not move because they saw no **parallax,** the apparent motion of an object because of the motion of the observer. To demonstrate parallax, close one eye and cover a distant object with your thumb held at arm's length. Switch eyes, and your thumb appears to shift position as shown at left. If Earth moves, ancient astronomers reasoned, you should see the sky from different locations at different times of the year, and you should see parallax distorting the shapes of the constellations. They saw no parallax, so they concluded Earth could not move. Actually, the parallax of the stars is too small to see with the unaided eye.

Every 2.14 years, Mars passes through a retrograde loop. Two successive loops are shown here. Each loop occurs further east along the ecliptic and has its own shape.

2 Planetary motion was a big problem for ancient astronomers. In fact, the word *planet* comes from the Greek word for "wanderer," referring to the eastward motion of the planets against the background of the fixed stars. The planets did not, however, move at a constant rate, and they could occasionally stop and move westward for a few months before resuming their eastward motion. This backward motion is called **retrograde motion.**

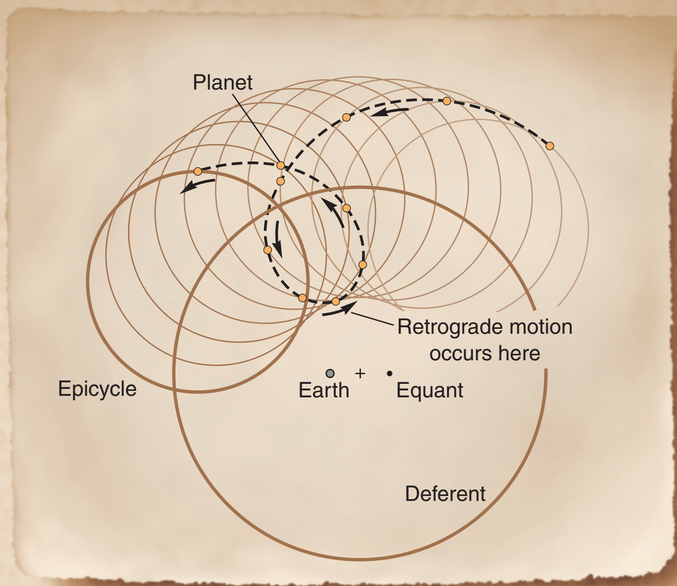


2a Simple uniform circular motion centered on Earth could not explain retrograde motion, so ancient astronomers combined uniformly rotating circles much like gears in a machine to try to reproduce the motion of the planets.

3 Uniformly rotating circles were key elements of ancient astronomy. Claudius Ptolemy created a mathematical model of the Aristotelian universe in which the planet followed a small circle called the **epicycle** that slid around a larger circle called the **deferent.** By adjusting the size and rate of rotation of the circles, he could approximate the retrograde motion of a planet. See illustration at right.

To adjust the speed of the planet, Ptolemy supposed that Earth was slightly off center and that the center of the epicycle moved such that it appeared to move at a constant rate as seen from the point called the **equant.**

To further adjust his model, Ptolemy added small epicycles (not shown here) riding on top of larger epicycles, producing a highly complex model.

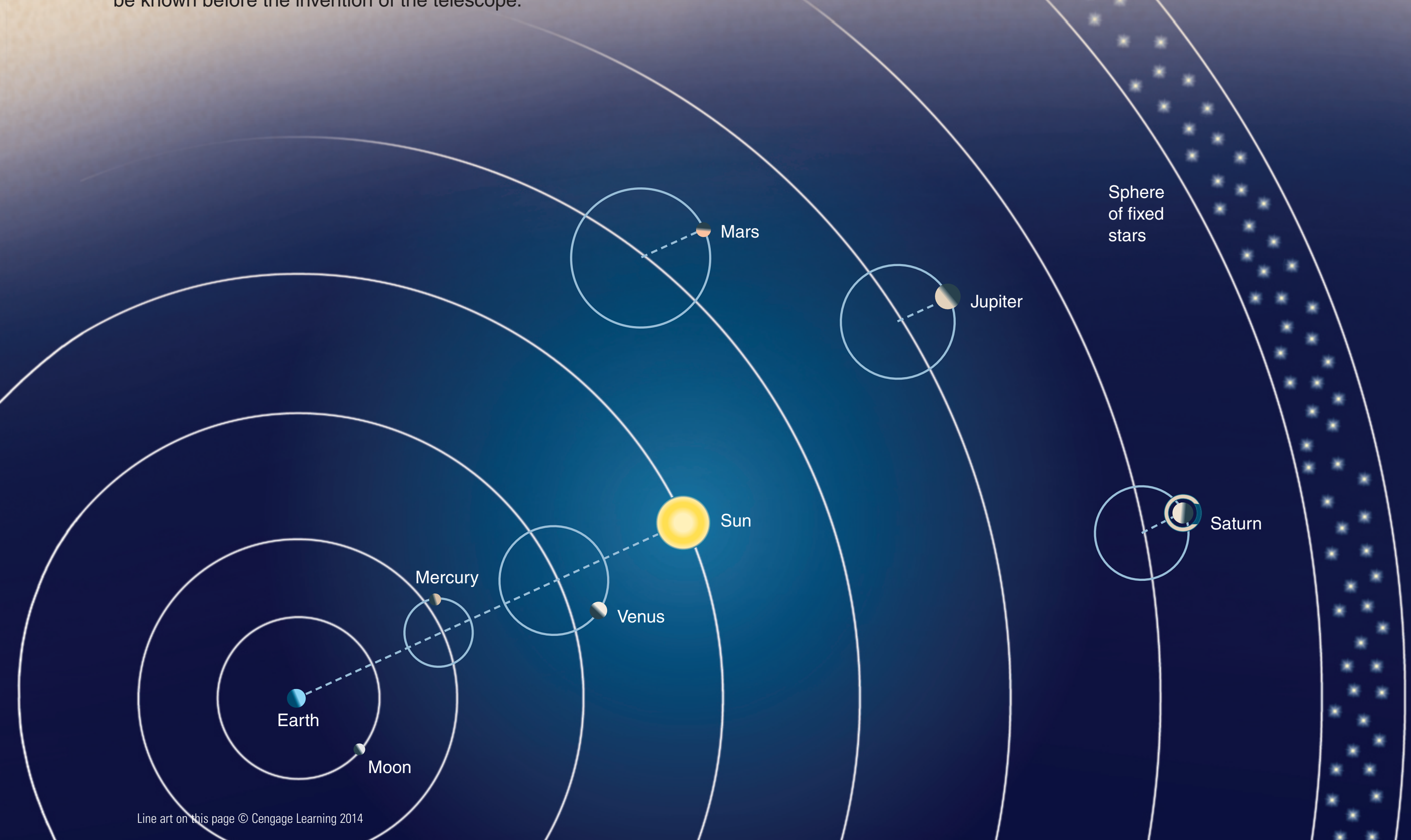


3a Ptolemy's great book *Mathematical Syntaxis* (c. AD 140) contained the details of his model. Islamic astronomers preserved and studied the book through the Middle Ages, and they called it *Al Magisti* (The Greatest). When the book was found and translated from Arabic to Latin in the 12th century, it became known as *Almagest*.

3b The Ptolemaic model of the universe shown below was geocentric and based on uniform circular motion. Note that Mercury and Venus were treated differently from the rest of the planets. The centers of the epicycles of Mercury and Venus had to remain on the Earth–Sun line as the sun circled Earth through the year.

Equants and smaller epicycles are not shown here. Some versions contained nearly 100 epicycles as generations of astronomers tried to fine-tune the model to better reproduce the motion of the planets.

Notice that this modern illustration shows rings around Saturn and sunlight illuminating the globes of the planets, features that could not be known before the invention of the telescope.



Line art on this page © Cengage Learning 2014

The Alfonsine Tables, the last great attempt to make the Ptolemaic system of practical use.

In Chapter 1, the cosmic zoom gave you a preview of the scale of the universe as you expanded your field of view from Earth to include our solar system, our galaxy, and finally billions of other galaxies. The ancients didn't know about stars and galaxies, and they imagined a universe that was much smaller. Earth lay at the center of their universe surrounded by crystalline shells carrying the planets with the starry sphere just beyond the outermost planetary shell.

Scholars and educated people knew Aristotle's astronomy well. You may have heard the **Common Misconception** that Christopher Columbus had to convince Queen Isabella of Spain that the world was round and not flat. Not so. Like all educated people of her time, the queen knew the world was round. Aristotle said so. Columbus had to convince the queen that the world was *small*—so small he could sail to the Orient by heading west. In making his sales pitch, he underestimated the size of Earth and overestimated the eastward extent of Asia, so he thought China and Japan were only a modest voyage west of Spain. If North America had not been in his way, he and his crew would have starved to death long before they reached Japan.

SCIENTIFIC ARGUMENT

Why did classical astronomers conclude the heavens were made up of spheres?

Today, scientific arguments depend on evidence and theory, but, in classical times, philosophers reasoned from first principles. Plato argued that the perfect geometrical figure was a sphere. Then the heavens, which everyone agreed were perfect, must be made up of spheres. The natural motion of a sphere is rotation, and the only perfect motion is uniform motion, so the heavenly spheres were thought to move in uniform circular motion. In this way, classical philosophers argued that the daily motion of the heavens around Earth and the motions of the seven planets (the sun and moon were counted as planets) against the background of the stars had to be produced by the combination of uniformly rotating spheres carrying objects around in perfect circles.

Now build a new argument. Although ancient astronomers didn't use evidence as modern scientists do, they did observe the world around them. **What observations led them to conclude that Earth didn't move?**

4-2 The Copernican Revolution

You would not have expected Nicolaus Copernicus to trigger a revolution in astronomy and science. He was born in 1473 to a merchant family in Poland. Orphaned at the age of 10, he was raised by his uncle, an important bishop, who sent him to the University of Kraców and then to the best universities in Italy where he studied law and medicine. Nevertheless, he had a passion for astronomy even as a student (■ Figure 4-1).

The Copernican Model

If you had sat beside Copernicus in his astronomy classes, you would have studied the Ptolemaic universe. The central location of Earth was widely accepted, and everyone knew that the heavens moved in uniform circular motion. For most scholars, questioning these principles was not an option, because, over the course of centuries, Aristotle's proposed geometry had become linked with Christian teachings. According to the Aristotelian universe, the most perfect region was the starry sphere and the most imperfect was Earth's center. This classical geocentric universe matched the commonly held Christian geometry of heaven and hell, so anyone who criticized the Ptolemaic model was not only questioning Aristotle's geometry but also indirectly challenging belief in heaven and hell.

For this reason, Copernicus probably found it difficult at first to consider alternatives to the Ptolemaic universe. Throughout his life, he was associated with the Catholic Church, which had adopted many of Aristotle's ideas. His uncle was an important bishop in Poland, and, through his uncle's influence, Copernicus was appointed a canon at the cathedral in Frauenberg at the unusually young age of 24. (A canon was not a priest but a Church administrator.) This gave Copernicus an income, although he continued his studies at the universities in Italy. When he left the universities, he joined his uncle and served as his secretary and personal physician until his uncle died in 1512. At that point, Copernicus moved into quarters adjoining the cathedral in Frauenburg, where he served as canon for the rest of his life.

His close connection with the Church notwithstanding, Copernicus began to consider an alternative to the Ptolemaic universe, probably while he was still at university. Sometime before 1514, he wrote an essay proposing a model of a **heliocentric universe** in which the sun, not Earth, was the center. To explain the daily and annual cycles of the sky, he proposed that Earth rotates on its axis and revolves around the sun. He distributed this commentary in handwritten form, without a title, and in some cases anonymously, to friends and astronomical correspondents. He may have been cautious out of modesty, or out of respect for the Church, or out of fear that his revolutionary ideas would be attacked unfairly. After all, the place of Earth was a controversial theological subject. Although this early essay discusses every major aspect of his later work, it did not include observations and calculations. His ideas needed support, so he began gathering observations and making detailed calculations that he planned to publish as a book that would demonstrate the truth of his revolutionary idea.

De Revolutionibus

Copernicus worked on his book *De Revolutionibus Orbium Coelestium (The Revolutions of the Celestial Spheres)* over a period

Fair Use

■ **Figure 4-1**

Nicolaus Copernicus (1473–1543) pursued a lifetime career in the Church, but he was also a talented mathematician and astronomer. His work triggered a revolution in human thought. These stamps were issued in 1973 to mark the 500th anniversary of his birth.

of many years and was essentially finished by about 1529; yet he hesitated to publish it even though other astronomers already knew of his theories. Even Church officials, concerned about the reform of the calendar, sought his advice and looked forward to the publication of his book.

One reason he hesitated was that he knew that the idea of a heliocentric universe would be highly controversial. This was a time of rebellion in the Church—Martin Luther (1483–1546) was speaking harshly about fundamental Church teachings, and others, both scholars and scoundrels, were questioning the Church's authority. Even matters as abstract as astronomy could stir controversy. Remember, too, that Earth's place in astronomical theory was linked to the geometry of heaven and hell, so moving Earth from its central place was a controversial and perhaps heretical idea.

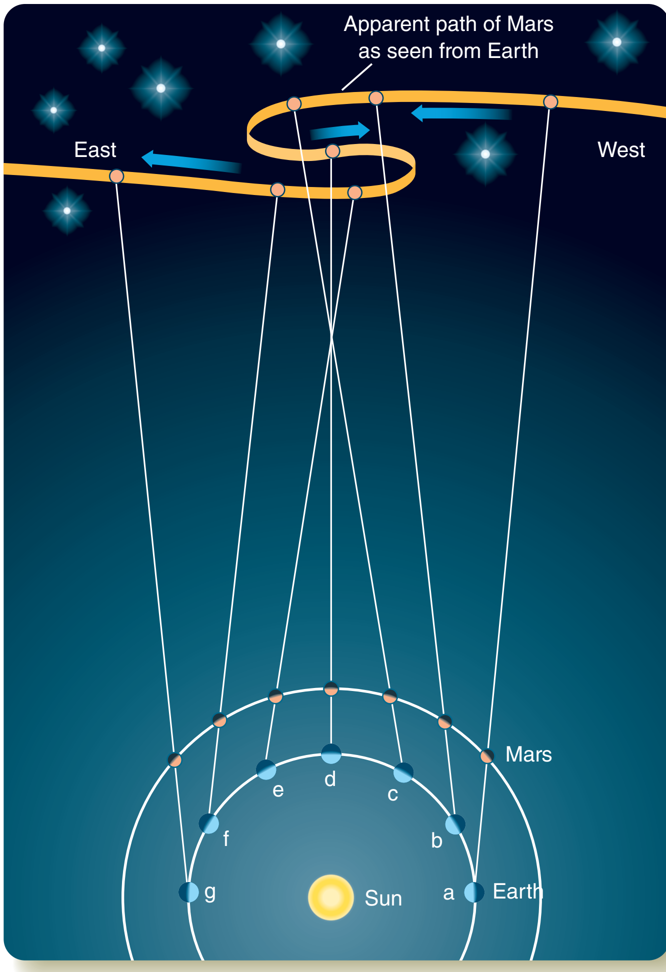
Another reason Copernicus may have hesitated to publish was that his work was incomplete. His model could not accurately predict planetary positions, so he continued to refine it. Finally in 1540 he allowed the visiting astronomer Joachim Rheticus (1514–1576) to publish an account of the Copernican universe in Rheticus's book *Narratio Prima (First Narrative).* In 1542, Copernicus sent the manuscript for *De Revolutionibus* off to be printed. He died in the spring of 1543 before the printing was completed.

The most important idea in the book was the location of the sun at the center of the universe. That single innovation had an astonishing consequence—the retrograde motion of the planets was immediately explained in a straightforward way without the large epicycles that Ptolemy had used.

In the Copernican system, Earth moves faster along its orbit than the planets that lie farther from the sun. Consequently, Earth periodically overtakes and passes these planets. To visualize this, imagine that you are in a race car, driving rapidly along the inside lane of a circular racetrack. As you pass slower cars driving in the outer lanes, they fall behind, and if you did not realize you were moving, it would look as if the cars in the outer lanes occasionally slowed to a stop and then backed up for a short interval. ■ Figure 4-2 shows how the same thing happens as Earth passes a planet such as Mars. Although Mars moves steadily along its orbit, as seen from Earth it appears to slow to a stop and move westward (retrograde) as Earth passes it. This happens to any planet whose orbit lies outside Earth's orbit, so Mars, Jupiter, and Saturn occasionally move retrograde along the ecliptic. Because the planetary orbits do not lie in precisely the same plane, a planet does not resume its eastward motion in precisely the same path it followed earlier. Consequently, it describes a loop whose shape depends on the angle between the orbital planes and the planet's location along the ecliptic.

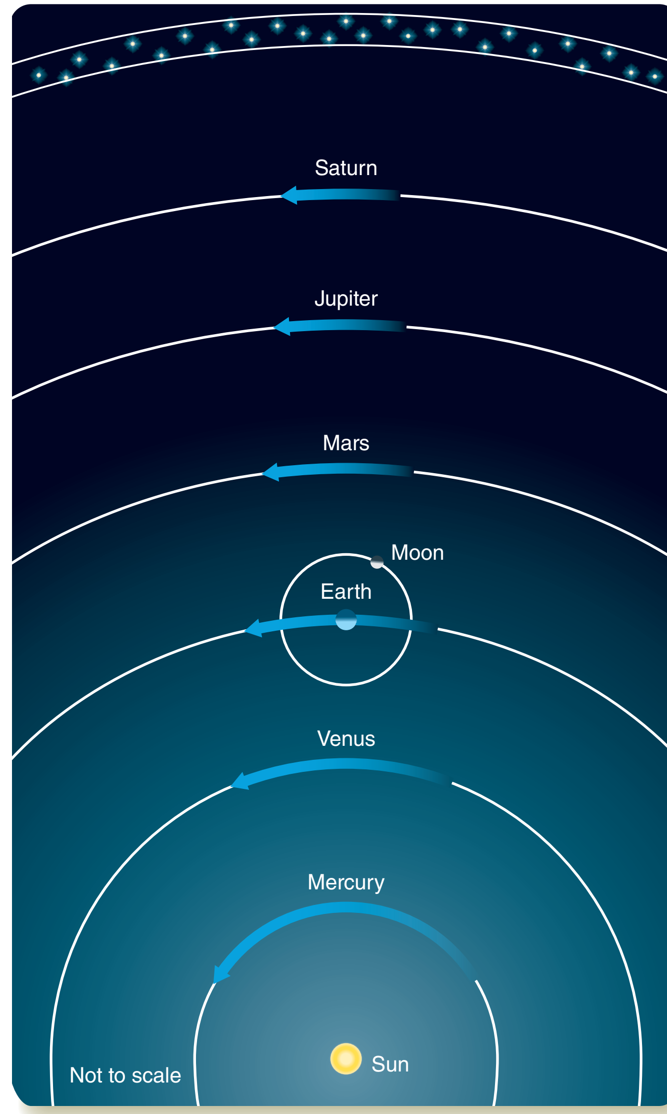
Copernicus could explain retrograde motion without epicycles, and that was impressive. The Copernican system was elegant and simple compared with the whirling epicycles and off-center equants of the Ptolemaic system. You can see Copernicus's own diagram for his heliocentric system in the top stamp in Figure 4-1. However, *De Revolutionibus* failed in one critical way—the Copernican model could not predict the positions of the planets significantly more accurately than the Ptolemaic system could. To understand why it failed, you must understand Copernicus and his world.

Copernicus proposed a revolutionary idea when he made the planetary system heliocentric, but he was a classical astronomer with tremendous respect for the old concept of uniform circular motion. In fact, Copernicus objected strongly to Ptolemy's use of the equant because it required that the planet move faster in some places and slower in others. It seemed arbitrary to Copernicus, an obvious violation of uniform circular motion and the elegance of Aristotle's philosophy of the heavens. Copernicus called equants



■ **Figure 4-2**

The Copernican explanation of retrograde motion. As Earth overtakes Mars (a–c), Mars appears to slow its eastward motion. As Earth passes Mars (d), Mars appears to move westward. As Earth draws ahead of Mars (e–g), Mars resumes its eastward motion against the background stars. The positions of Earth and Mars are shown at equal intervals of 1 month.



■ **Figure 4-3**

The Copernican universe was elegant in its arrangement and its motions. Mercury and Venus are treated just like all the other planets, and orbital velocities (blue arrows) decrease smoothly from that of Mercury, the fastest, to that of Saturn, the slowest. Compare the elegance of this model with the complexity of the Ptolemaic model as shown on page 49.

"monstrous" because they undermined both geocentrism and uniform circular motion. In devising his model, Copernicus demonstrated a strong belief in uniform circular motion.

Although he did not need epicycles to explain retrograde motion, Copernicus quickly discovered that the sun, moon, and planets suffered other small variations in their motions that he could not explain using uniform circular motion centered on the sun. Today astronomers recognize that those variations are the result of planets following slightly noncircular elliptical orbits, but because Copernicus held firmly to uniform circular motion, he had to introduce small epicycles to try to reproduce these minor variations in the motions of the sun, moon, and planets.

Because Copernicus imposed uniform circular motion on his model, it could not accurately predict the motions of the planets. *The Prutenic Tables* (1551) were based on the Copernican model, and they were not significantly more accurate than the 13th-century *Alfonsine Tables* that were based on Ptolemy's model. Both could be in error by as much as 2°, which is four times the angular diameter of the full moon.

The Copernican *model* is inaccurate. It includes uniform circular motion and consequently does not precisely describe the motions of the planets. But the Copernican *hypothesis* that the universe is heliocentric is correct. Considering how little astronomers of the time knew of other stars and galaxies, the universe that Copernicus knew was heliocentric in that the planets circle the sun and not Earth.

Although astronomers throughout Europe read and admired *De Revolutionibus,* they did not immediately accept the

How Do We Know? 4-1

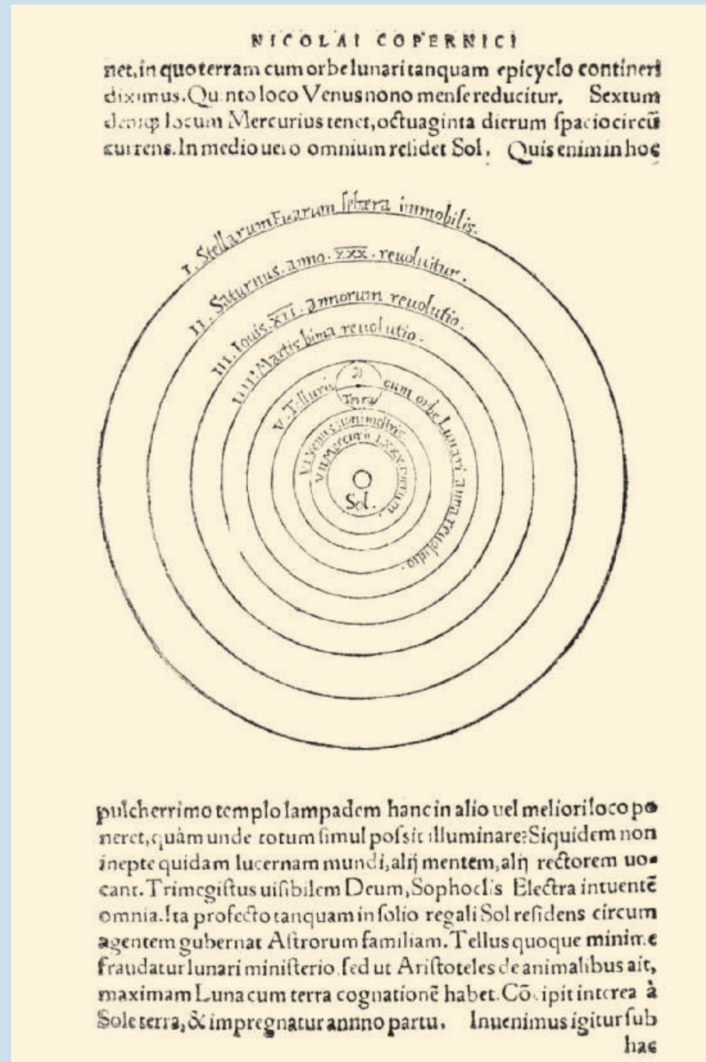
Scientific Revolutions

How do scientific revolutions occur? You might think from what you know of the scientific method that science grinds forward steadily as new theories are tested against evidence and accepted or rejected. In fact, science sometimes leaps forward in scientific revolutions. The Copernican Revolution is often cited as the perfect example; in a few decades, astronomers rejected the 2000-year-old geocentric model and adopted the heliocentric model. Why does that happen? It's all because scientists are human.

The American philosopher of science Thomas Kuhn has referred to a commonly accepted set of scientific ideas and assumptions as a scientific **paradigm.** The pre-Copernican astronomers shared a geocentric paradigm that included uniform circular motion, geocentrism, and the perfection of the heavens. Although they were intelligent, they were prisoners of that paradigm. A scientific paradigm is powerful because it shapes your perceptions. It determines what you judge to be important questions and what you judge to be significant evidence. Consequently, the ancient astronomers could not recognize how their geocentric paradigms limited what they understood.

You will see here how the work of Copernicus, Galileo, and Kepler overthrew the geocentric paradigm. Scientific revolutions occur when the deficiencies of the old paradigm build up until finally a scientist has the insight to think "outside the box." Pointing out the failings of the old ideas and proposing a new paradigm with supporting evidence is like poking a hole in a dam; suddenly the pressure is released, and the old paradigm is swept away.

Scientific revolutions are exciting because they give you a dramatic new understanding of nature, but they are also times of conflict as new insights sweep away old ideas.



NICOLAI COPERNICI

net, in quo terram cum orbe lunari tanquam epicyclo contineri diximus. Quinto loco Venus nono mense reducitur. Sextum denique locum Mercurius tenet, octuaginta dierum spacio circu currens. In medio uero omnium residet Sol. Quis enim in hoc

pulcherrimo templo lampadem hanc in alio uel meliori loco poneret, quàm unde totum simul possit illuminare? Siquidem non inepte quidam lucernam mundi, alij mentem, alij rectorem uocant. Trimegistus uisibilem Deum, Sophoclis Electra intuentē omnia. Ita profecto tanquam in solio regali Sol residens circum agentem gubernat Astrorum familiam. Tellus quoque minime fraudatur lunari ministerio, sed ut Aristoteles de animalibus ait, maximam Luna cum terra cognationē habet. Cōcipit interea à Sole terra, & impregnatur annno partu. Inuenimus igitur sub hac

Yerkes Observatory

The Copernican universe as shown in De Revolutionibus changed science.

Copernican hypothesis. The mathematics was elegant, and the astronomical observations and calculations were of tremendous value, but few astronomers believed, at first, that the sun actually was the center of the planetary system and that Earth moved. How the Copernican hypothesis was gradually recognized as correct has been called the Copernican Revolution because it was not just the adoption of a new idea but a total change in the way astronomers, and, in fact, all of humanity, thought about the place of the Earth (**How Do We Know? 4-1**).

There are probably a number of reasons why the Copernican hypothesis gradually won support, including the revolutionary temper of the times, but the most important factor may have been the elegance of the idea. Placing the sun at the center of the universe produced a symmetry among the motions of the planets that is pleasing to the eye as well as to the intellect (■ Figure 4-3). In the Ptolemaic model, Mercury and Venus were treated differently from the rest of the planets; their epicycles had to remain centered on the Earth–sun line. In the Copernican model, all of the planets were treated the same. They all followed orbits that circled the sun at the center. Furthermore, their speed depended in an orderly way on their distance from the sun, with those closest moving fastest.

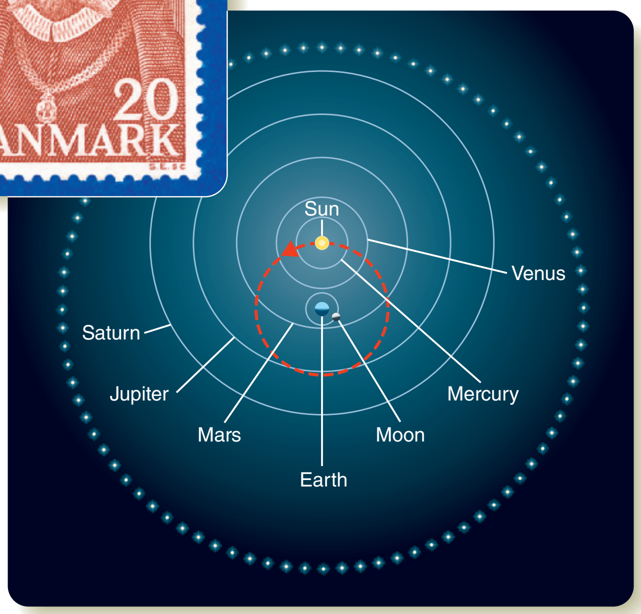


Figure 4-4

Tycho Brahe (1546–1601) was, during his lifetime, the most famous astronomer in the world. Proud of his noble rank, he wears the elephant medal awarded him by the king of Denmark. His artificial nose is suggested in this engraving. Tycho Brahe's model of the universe retained the first principles of classical astronomy; it was geocentric with the sun and moon revolving around Earth, and the planets revolving around the sun. All motion was along circular paths.

The most astonishing consequence of the Copernican hypothesis was not what it said about the sun but what it said about Earth. By placing the sun at the center, Copernicus made Earth into a planet, moving along an orbit like the other planets. By making Earth a planet, Copernicus revolutionized humanity's view of its place in the universe and triggered a controversy that would eventually bring the astronomer Galileo Galilei before the Inquisition. This controversy over the apparent conflict between scientific knowledge and philosophical and theological beliefs continues even today.

SCIENTIFIC ARGUMENT

Why would you say the Copernican hypothesis was correct but the model was inaccurate?

To build this argument, you must distinguish carefully between a hypothesis and a model. The Copernican hypothesis was that the sun and not Earth was the center of the universe. Given the limited knowledge of the Renaissance astronomers about distant stars and galaxies, that hypothesis was correct.

The Copernican model, however, included not only the heliocentric hypothesis but also uniform circular motion. The model is inaccurate because the planets don't really follow circular orbits, and the small epicycles that Copernicus added to his model never quite reproduced the motions of the planets.

Now build a new argument. The Copernican hypothesis won converts because it is elegant and can explain retrograde motion. **How does its explanation of retrograde motion work, and how is it more elegant than the Ptolemaic explanation?**

4-3 Planetary Motion

The Copernican Hypotheisis solved the problem of the place of Earth, but it didn't explain planetary motion. If planets don't move in uniform circular motion, how do they move? The puzzle of planetary motion was solved during the century following the death of Copernicus through the work of two men. One compiled the observations, and the other did the analysis.

Tycho Brahe

Tycho Brahe (1546–1601) was not a churchman like Copernicus but rather a nobleman from an important family, educated at the finest universities. He was well known for his vanity and his lordly manners, and by all accounts he was a proud and haughty nobleman. Tycho's disposition was not improved by a dueling injury from his university days. His nose was badly disfigured, and for the rest of his life he wore false noses made of gold and silver, stuck on with wax (■ Figure 4-4).

Although Tycho officially studied law at the university, his real passions were mathematics and astronomy, and early in his university days he began measuring the positions of the planets in the sky. In 1563, Jupiter and Saturn passed very near each other in the sky, nearly merging into a single point on the night of August 24. Tycho found that the *Alfonsine Tables* were a full month in error and that the *Prutenic Tables* were in error by a number of days.

In 1572, a "new star" (now called Tycho's supernova) appeared in the sky, shining more brightly than Venus, and Tycho carefully measured its position. According to classical astronomy, the new star represented a change in the heavens and therefore had to lie below the sphere of the moon. Tycho, even though he believed in a geocentric universe, understood that the new star should show parallax, meaning that it would appear slightly too far east as it rose and slightly too far west as it set. But Tycho saw no parallax in the position of the new star, so he concluded that it must lie above the sphere of the moon and was probably on the starry sphere itself. This contradicted Aristotle's conception of the starry sphere as perfect and unchanging.

No one before Tycho could have made this discovery because no one had ever measured the positions of celestial objects so accurately. Tycho had great confidence in the precision of his measurements, and he had studied astronomy thoroughly, so when he failed to detect parallax for the new star, he knew it was important evidence against the Ptolemaic theory. He announced his discovery in a small book, *De Stella Nova (The New Star)*, published in 1573.

The book attracted the attention of astronomers throughout Europe, and soon Tycho's family introduced him to the court of the Danish king Frederick II, where he was offered funds to build an observatory on the island of Hveen just off the Danish coast. To support his observatory, Tycho was given a steady income as lord of a coastal district from which he collected rents. (He was not a popular landlord.) On Hveen, Tycho constructed a luxurious home with six towers especially equipped for astronomy and populated it with servants, assistants, and a dwarf to act as jester. Soon Hveen was an international center of astronomical study.

Tycho Brahe's Legacy

Tycho made no direct contribution to astronomical theory. Because he could measure no parallax for the stars, he concluded that Earth had to be stationary, thus rejecting the Copernican hypothesis. However, he also rejected the Ptolemaic model because of its inaccuracy. Instead he devised a complex model in which Earth was the immobile center of the universe around which the sun and moon moved. The other planets circled the sun (Figure 4-4). The model thus incorporated part of the Copernican model, but in it Earth—not the sun—was stationary. In this way, Tycho preserved the central immobile Earth. Although Tycho's model was very popular at first, the Copernican model replaced it within a century.

The true value of Tycho's work was observational. Because he was able to devise new and better instruments, he was able to make highly accurate observations of the position of the stars, sun, moon, and planets. Tycho had no telescopes—they were not invented until the next century—so his observations were made by the unaided eye peering through peepholes much like gun sights. He and his assistants made precise observations for 20 years at Hveen.

Unhappily for Tycho, King Fredrick II died in 1588, and his young son took the throne. Suddenly, Tycho's temper, vanity, and noble presumptions threw him out of favor. In 1596, taking most of his instruments and books of observations, he went to Prague, the capital of Bohemia, and became imperial mathematician to the Holy Roman Emperor Rudolph II. His goal was to revise the *Alfonsine Tables* and publish the result as a monument to his new patron. It would be called the *Rudolphine Tables.*

Tycho did not intend to base the *Rudolphine Tables* on the Ptolemaic system but rather on his own Tyconic system, proving once and for all the validity of his hypothesis. To assist him, he hired a few mathematicians and astronomers, including one Johannes Kepler. Then, in November 1601, Tycho collapsed at a nobleman's home. Before he died, 11 days later, he asked Rudolph II to make Kepler imperial mathematician. The newcomer became Tycho's replacement (though at one-sixth Tycho's salary).

Kepler: An Astronomer of Humble Origins

No one could have been more different from Tycho Brahe than Johannes Kepler (■ Figure 4-5). Kepler was born in 1571 to a poor family in a region that is now part of southwest Germany. His father was unreliable and shiftless, principally employed as a mercenary soldier fighting for whoever paid enough. He was often absent for long periods and finally failed to return from a military expedition. Kepler's mother was apparently an unpleasant and unpopular woman. She was accused of witchcraft in later years, and Kepler had to defend her in a trial that dragged on for three years. She was finally acquitted, but she died the following year.

In spite of family disadvantages and chronic poor health, Kepler did well in school, winning promotion to a Latin school and eventually a scholarship to the university at Tübingen, where he studied to become a Lutheran pastor. During his last year of study, Kepler accepted a job in Graz teaching mathematics and astronomy, a job he resented because, as he claimed in a letter to a friend, he knew little about the subjects. Evidently he was not a good teacher—he had few students his first year and none at all his second. His superiors put him to work teaching a few introductory courses and preparing an annual almanac that contained astronomical, astrological, and weather predictions. Through good luck, in 1595 some of his weather predictions were fulfilled, and he gained a reputation as an astrologer and seer. Even in later life he earned money from his almanacs.

While still a college student, Kepler had adopted the Copernican hypothesis, and at Graz he used his extensive spare time to study astronomy. By 1596, the same year Tycho arrived in Prague, Kepler was sure he had solved the mystery of the universe. That year he published a book called *The Forerunner of Dissertations on the Universe, Containing the Mystery of the Universe.* The book, like nearly all scientific works of that age, was written in Latin and is now known as *Mysterium Cosmographicum.*

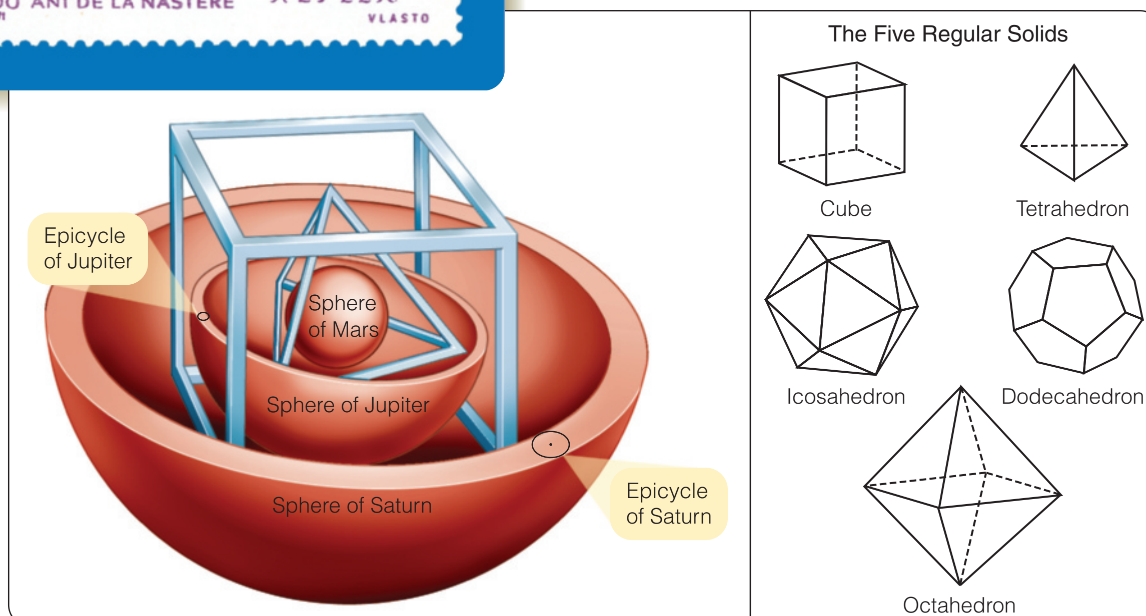
By modern standards, the book contains almost nothing of value. It begins with a long appreciation of Copernicanism and then goes on to speculate on the reasons for the spacing of the planetary orbits. Kepler assumed that the heavens could be described by only the most perfect of shapes. Therefore he felt that he had found the underlying architecture of the universe in the sphere plus the five regular solids.* In Kepler's model, the five regular solids became spacers for the orbits of the six planets, which were represented by nested spheres (Figure 4-5). In fact, Kepler concluded that there could be only six planets (Mercury, Venus, Earth, Mars, Jupiter, and Saturn) because there were only five regular solids to act as spacers between their spheres. He provided astrological, numerological, and even musical arguments for his theory.

The second half of the book is no better than the first, but it has one virtue—as Kepler tried to fit the five solids to the planetary orbits, he demonstrated that he was a talented mathematician and that he was well versed in astronomy. He sent copies of his book to Tycho on Hveen and to Galileo in Rome.

*The five regular solids, also known as the Platonic solids, are the tetrahedron, cube, octahedron, dodecahedron, and icosahedron. They were considered perfect because the faces and the angles between the faces are the same at every corner.

Figure 4-5

Johannes Kepler (1571–1630) was Tycho Brahe's successor. This diagram, based on one drawn by Kepler, shows how he believed the sizes of the celestial spheres carrying the outer three planets—Saturn, Jupiter, and Mars—are determined by spacers (blue) consisting of two of the five regular solids. Inside the sphere of Mars, the remaining regular solids (not shown in this drawing) separate the spheres of Earth, Venus, and Mercury. The sun lay at the very center of this Copernican universe based on geometrical spacers.



Joining Tycho

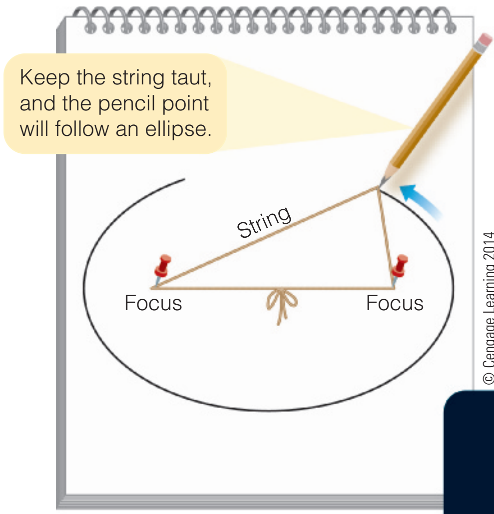
Life was unsettled for Kepler because of the persecution of Protestants in the region, so when Tycho Brahe invited him to Prague in 1600, Kepler went readily, eager to work with the famous Danish astronomer. Tycho's sudden death in 1601 left Kepler, the new imperial mathematician, in a position to use the observations from Hveen to analyze the motions of the planets and complete *The Rudolphine Tables.* Tycho's family, recognizing that Kepler was a Copernican and guessing that he would not follow the Tychonic system in completing *The Rudolphine Tables,* sued to recover the instruments and books of observations. The legal wrangle went on for years. Tycho's family did get back the instruments Tycho had brought to Prague, but Kepler had the books, and he kept them.

Whether Kepler had any legal right to Tycho's records is debatable, but he put them to good use. He began by studying the motion of Mars, trying to deduce from the observations how the planet moved. By 1606, he had solved the mystery, this time correctly. The orbit of Mars is not a circle but a very slightly elongated ellipse, and with that he abandoned the 2000-year-old belief in the circular motion of the planets. But even this insight was not enough to explain the observations. The planets do not move at uniform speeds along their elliptical orbits. Kepler's analysis showed that they move faster when close to the sun and slower when farther away. With those two brilliant discoveries, Kepler abandoned uniform circular motion and finally solved the puzzle of planetary motion. He published his results in 1609 in a book called *Astronomia Nova (New Astronomy).*

In spite of the abdication of Rudolph II in 1611, Kepler continued his astronomical studies. He wrote about a supernova that appeared in 1604 (now known as Kepler's supernova) and about comets, and he wrote a textbook about Copernican astronomy. In 1619, he published *Harmonices Mundi (The Harmony of the World),* in which he returned to the cosmic mysteries of *Mysterium Cosmographicum.* The only thing of note in *Harmonices Mundi* is his discovery that the radii of the planetary orbits are related to the planets' orbital periods. That and his two previous discoveries are so important that they have become known as the three most fundamental rules of orbital motion.

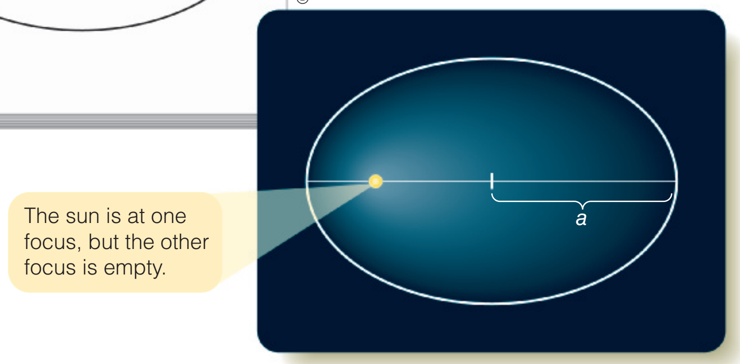
Kepler's Three Laws of Planetary Motion

Although Kepler dabbled in the philosophical arguments of his day, he was at heart a mathematician, and his triumph was his explanation of the motion of the planets. The key to his solution was the ellipse.



■ **Figure 4-6**

The geometry of elliptical orbits: Drawing an ellipse with two tacks and a loop of string is easy. The semimajor axis, *a,* is half of the longest diameter. The sun lies at one of the foci of the elliptical orbit of a planet.



An **ellipse** is a figure that can be drawn around two points, called the *foci,* in such a way that the distance from one focus to any point on the ellipse and back to the other focus equals a constant. This makes it easy to draw ellipses using two thumbtacks and a loop of string. Press the thumbtacks into a board, loop the string about the tacks, and place a pencil in the loop. If you keep the string taut as you move the pencil, it traces out an ellipse (■ Figure 4-6).

The geometry of an ellipse is described by two simple numbers. The **semimajor axis, *a***, is half of the longest diameter, as you can see in Figure 4-6. The **eccentricity, *e***, of an ellipse is half the distance between the foci divided by the semimajor axis. The eccentricity of an ellipse tells you its shape; if *e* is nearly equal to one, the ellipse is very elongated. If *e* is closer to zero, the ellipse is more circular. To draw a circle with the string and tacks shown Figure 4-6, you would have to move the two thumbtacks together because a circle is really just an ellipse with eccentricity equal to zero. Try fiddling with real thumbtacks and string, and you'll be surprised how easy it is to draw graceful, smooth ellipses with various eccentricities.

Ellipses are a prominent part of Kepler's three fundamental rules of planetary motion. Those rules have been tested and confirmed so many times that astronomers now refer to them as natural laws (**How Do We Know? 4-2**). They are commonly called Kepler's laws of planetary motion (■ Table 4-1).

Kepler's first law says that the orbits of the planets around the sun are ellipses with the sun at one focus. Thanks to the precision of Tycho's observations and the sophistication of Kepler's mathematics, Kepler was able to recognize the very slightly elongated elliptical shape of the orbits even though they are nearly circular. Mercury has the most elliptical orbit, but even it deviates only slightly from a circle (■ Figure 4-7).

Kepler's second law says that an imaginary line drawn from the planet to the sun always sweeps over equal areas in equal intervals of time. This means that when the planet is closer to the sun and the line connecting it to the sun is shorter, the planet moves more rapidly, and the line sweeps over the same area that is swept over when the planet is farther from the sun. You can see how the planet in Figure 4-7 would move from point *A* to point *B* in one month, sweeping over the area shown. But when the planet is farther from the sun, one month's motion would be shorter, from *A′* to *B′*, and the area swept out would be the same.

Kepler's third law relates a planet's orbital period to its average distance from the sun. The orbital period, *P,* is the time a planet takes to travel around the sun once. Its average distance from the sun turns out to equal the semimajor axis of its orbit, *a.* Kepler's third law says that a planet's orbital period squared is proportional to the semimajor axis of its orbit cubed. Measuring *P* in years and *a* in astronomical units, you can summarize the third law as

$$P_y^2 = a_{AU}^3$$

For example, Jupiter's average distance from the sun is roughly 5.2 AU. The semimajor axis cubed is about 140, so the period must be the square root of 140, which equals just under 12 years.

Notice that Kepler's three laws are empirical. That is, they describe a phenomenon without explaining why it occurs. Kepler derived the laws from Tycho's extensive observations, not from any first principle, fundamental assumption, or theory. In fact, Kepler never knew what held the planets in their orbits or why they continued to move around the sun.

The Rudolphine Tables

Kepler continued his mathematical work on *The Rudolphine Tables,* and at last, in 1627, it was ready. He financed the printing himself, dedicating the book to the memory of Tycho Brahe. In fact, Tycho's name appears in larger type on the title page than Kepler's own. This is surprising because the tables were not based on the Tyconic system but on the heliocentric model of Copernicus and the elliptical orbits of Kepler. The reason for Kepler's evident deference was Tycho's family, still powerful and still intent on protecting Tycho's reputation. They even demanded a share of the profits and the right to censor the book before publication, though they changed nothing but a few words on the title page and added an elaborate dedication to the emperor.

How Do We Know? 4-2

Hypothesis, Theory, and Law

Why is a theory much more than just a guess? Scientists study nature by devising and testing new hypotheses and then developing the successful ideas into theories and laws that describe how nature works.

A scientist's first step in solving a natural mystery is to propose a reasonable explanation based on what is known so far. This proposal, called a **hypothesis,** is a single assertion or statement that must be tested through observation and experimentation. A good illustration comes from the history of medicine. From the time of Aristotle philosophers believed that food spoils as a result of the spontaneous generation of life—for example, mold growing out of drying bread. French chemist Louis Pasteur (1822–1895) hypothesized that microorganisms were not spontaneously generated but were carried through the air. To test his hypothesis he sealed an uncontaminated nutrient broth in glass, completely protecting it from the microorganisms on dust particles in the air. No mold grew until he broke the seal and allowed air to contact the broth, effectively disproving spontaneous generation. Although others had argued against spontaneous generation before Pasteur, it was Pasteur's meticulous testing of his hypothesis through experimentation that finally convinced the scientific community.

A **theory** generalizes the specific results of well-confirmed hypotheses to give a broader description of nature, which can be applied to a wide variety of circumstances. For instance, Pasteur's specific hypothesis about mold growing in broth contributed to a broader theory that disease is caused by microorganisms transmitted from sick people to well people. This theory, called the germ theory of disease, is a cornerstone of modern medicine.

It is a **Common Misconception** that the word *theory* means a tentative idea, a guess. As you have just learned, scientists actually use the word *theory* to mean an idea that is widely applicable and confirmed by abundant evidence.

Sometimes, when a theory has been refined, tested, and confirmed so often that scientists have great confidence in it, it is called a **natural law.** Natural laws are the most fundamental principles of scientific knowledge. Kepler's laws of planetary motion are good examples.

Confidence is the key. In general, scientists have more confidence in a theory than in a hypothesis and the most confidence in a natural law. However, there is no precise distinction among a hypothesis, a theory, and a law, and use of these terms is sometimes a matter of tradition. For instance, some textbooks refer to the Copernican "theory" of heliocentrism, but it had not been well tested when Copernicus proposed it, and it is more rightly called the Copernican hypothesis. At the other extreme, Darwin's "theory" of evolution, containing many hypotheses that have been tested and confirmed over and over for nearly 150 years, might more correctly be called a natural law.

Collection of John Coolidge III/Franklin and Marshall College

A fossil of a 500-million-year-old trilobite: Darwin's theory of evolution has been tested many times and is universally accepted in the life sciences, but by custom it is called Darwin's theory and not Darwin's law.

Table 4-1 | Kepler's Laws of Planetary Motion

I. The orbits of the planets are ellipses with the sun at one focus.
II. A line from a planet to the sun sweeps over equal areas in equal intervals of time.
III. A planet's orbital period squared is proportional to its average distance from the sun cubed:

$$P_y^2 = a_{AU}^3$$

The Rudolphine Tables was Kepler's masterpiece. It could predict the positions of the planets 10 to 100 times more accurately than previous tables. Kepler's tables were the precise model of planetary motion that Copernicus had sought but failed to find. The accuracy of *The Rudolphine Tables* was strong evidence that both Kepler's laws of planetary motion and the Copernican hypothesis for the place of Earth were correct. Copernicus would have been pleased.

Kepler died in 1630. He had solved the problem of planetary motion, and his *Rudolphine Tables* demonstrated his solution. Although he did not understand why the planets moved or why they followed ellipses, insights that had to wait half a century for Isaac Newton, Kepler's three laws worked. In science the only test of a theory is, "Does it describe reality?" Kepler's laws have been used for almost four centuries as a true description of orbital motion.

SCIENTIFIC ARGUMENT

How was Kepler's model with regular solids based on first principles? How were his three laws based on evidence?

When he was younger, Kepler accepted Plato's argument for the perfection of the heavens. Furthermore, Kepler argued that the five regular solids were perfect geometrical figures and should be part of the perfect heavens along with spheres. He then arranged the five regular solids to produce the approximate spacing among the spheres that carried the planets in the Copernican model. Kepler's model was thus based on a first principle—the perfection of the heavens.

Much later, Kepler derived his three laws of motion from the observations made by Tycho Brahe during 20 years on Hveen. The observations were the evidence, and they gave Kepler a reality check each time he tried a new calculation. He chose ellipses because they fit the data and not because he thought ellipses had any special significance.

The Copernican model was a poor predictor of planetary motion, but *The Rudolphine Tables* was much more accurate. **What first principle did Copernicus follow that was abandoned when Kepler looked at the evidence?**

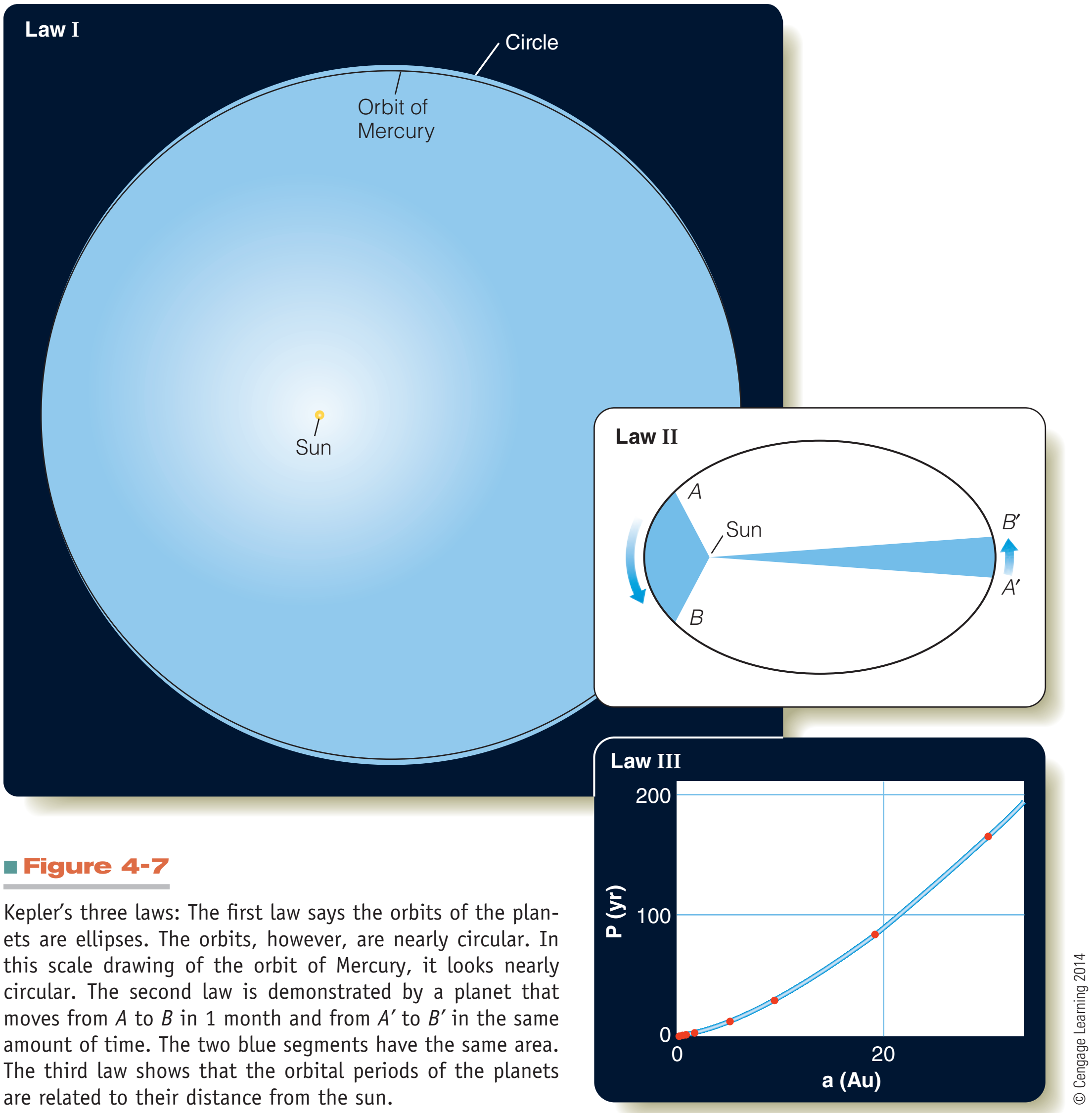


Figure 4-7

Kepler's three laws: The first law says the orbits of the planets are ellipses. The orbits, however, are nearly circular. In this scale drawing of the orbit of Mercury, it looks nearly circular. The second law is demonstrated by a planet that moves from *A* to *B* in 1 month and from *A′* to *B′* in the same amount of time. The two blue segments have the same area. The third law shows that the orbital periods of the planets are related to their distance from the sun.

4-4 Galileo Galilei

MOST PEOPLE THINK THEY KNOW two facts about Galileo, but both facts are wrong; they are **Common Misconceptions,** so you have probably heard them. Galileo did not invent the telescope, and he was not condemned by the Inquisition for believing that Earth moved around the sun. Then why is Galileo so famous? Why did the Vatican reopen his case in 1979, almost 400 years after his trial? As you learn about Galileo, you will discover that his trial concerned not just the place of Earth and the motion of the planets but also a new and powerful method of understanding nature, a method called science.

Telescopic Observations

Galileo Galilei (Figure 4-8) was born in 1564 in Pisa, a city in what is now Italy, and he studied medicine at the university there. His true love, however, was mathematics, and, although he had to leave school early for financial reasons, he returned only four years later as a professor of mathematics. Three years after that he became professor of mathematics at the university at Padua, where he remained for 18 years.

During this time, Galileo seems to have adopted the Copernican model, although he admitted in a 1597 letter to Kepler that he did not support Copernicanism publicly. At that time, the Copernican hypothesis was not officially considered heretical, but it was hotly debated among astronomers, and Galileo, living in a region controlled by the Church, cautiously avoided trouble. It was the telescope that finally drove Galileo to publicly defend the heliocentric model.

The telescope was apparently invented around 1608 by lens makers in Holland. Galileo, hearing descriptions in the fall of 1609, was able to build telescopes in his workshop. In fact, Galileo was not the first person to look at the sky through a telescope, but he was the first person to

Fair Use

■ **Figure 4-8**

Galileo Galilei (1564–1642), remembered as the great defender of Copernicanism, also made important discoveries in the physics of motion. He is honored here on an old Italian 2000-lira note.

apply telescopic observations to the theoretical problem of the day—the place of Earth.

What Galileo saw through his telescopes was so amazing that he rushed a small book into print. *Sidereus Nuncius (The Starry Messenger)* reported three major discoveries. First, the moon was not perfect. It had mountains and valleys on its surface, and Galileo even used some of the mountains' shadows to calculate their height. Aristotle's philosophy held that the moon was perfect, but Galileo showed that it was not only imperfect but was a world with features like Earth's.

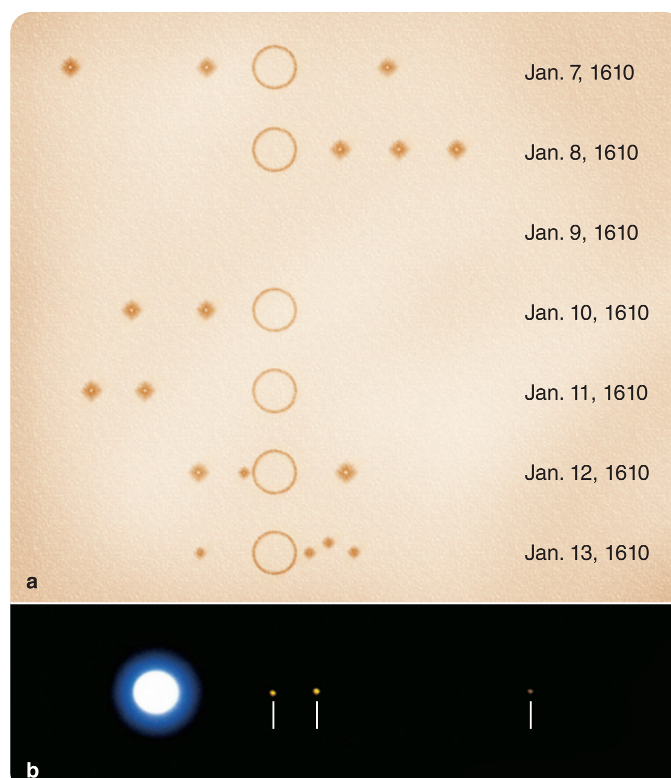
The second discovery reported in the book was that the Milky Way was made up of myriad stars too faint to see with the unaided eye. While intriguing, this could not match Galileo's third discovery. Galileo's telescope revealed four new "planets" circling Jupiter, objects known today as the **Galilean moons** of Jupiter (■ Figure 4-9).

The moons of Jupiter were strong evidence for the Copernican model. Critics of Copernicus had said Earth could not move because the moon would be left behind, but Galileo's discovery showed that Jupiter, which everyone agreed was moving, was able to keep its moons. That suggested that Earth, too, could move and keep its moon. Aristotle's philosophy also included the belief that all heavenly motion was centered on Earth. Galileo's observations showed that Jupiter's moons revolve around Jupiter, suggesting that there could be other centers of motion besides Earth.

Some time after *Sidereus Nuncius* was published, Galileo noticed something else that made Jupiter's moons even stronger evidence for the Copernican model. When he measured the orbital periods of the four moons, he found that the innermost moon moved fastest and that the moons further from Jupiter moved proportionally slower. Jupiter's moons made up a harmonious system ruled by Jupiter, just as the planets in the Copernican universe were a harmonious system ruled by the sun. (See Figure 4-3.) The similarity isn't proof, but Galileo saw it as an argument that the solar system is sun centered rather than Earth centered.

In the years following publication of *Sidereus Nuncius,* Galileo made two additional discoveries. When he observed the sun, he discovered sunspots, raising the suspicion that the sun was less than perfect. Further, by noting the movement of the spots, he concluded that the sun was a sphere and that it rotated on its axis.

His most dramatic discovery came when he observed Venus. Galileo saw that it was going through phases like those of the moon. In the Ptolemaic model, Venus moves around an epicycle



© Cengage Learning 2014

Grundy Observatory

■ **Figure 4-9**

(a) On the night of January 7, 1610, Galileo saw three small "stars" near the bright disk of Jupiter and sketched them in his notebook. On subsequent nights (excepting January 9, which was cloudy), he saw that the stars were actually four moons orbiting Jupiter. (b) This photograph taken through a modern telescope shows the overexposed disk of Jupiter and three of the four Galilean moons.

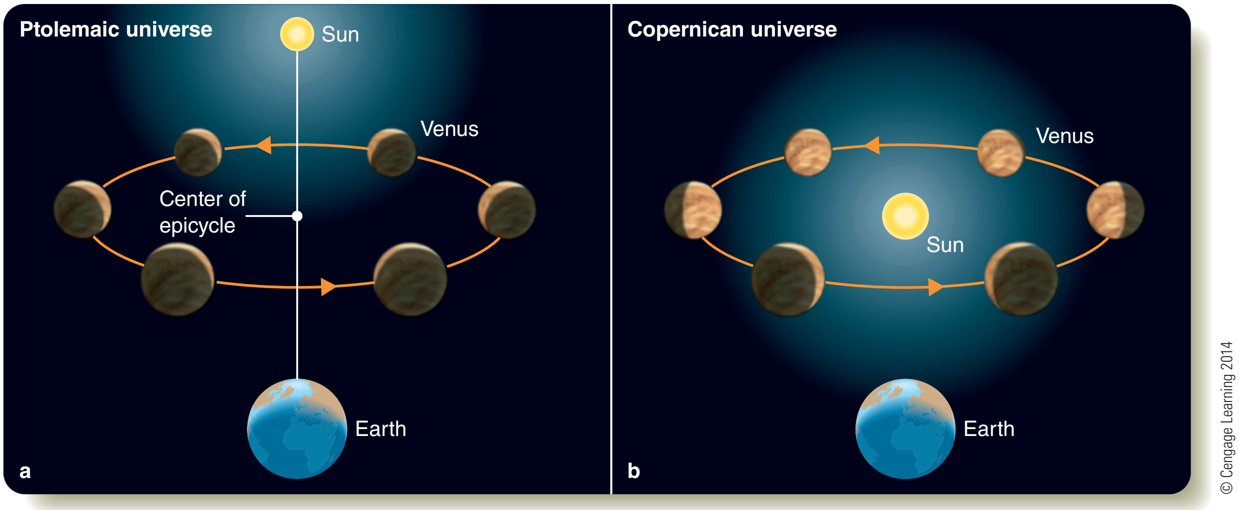


Figure 4-10

(a) If Venus moved in an epicycle centered on the Earth–sun line (see page 49), it would always appear as a crescent. (b) Galileo's telescope showed that Venus goes through a full set of phases, proving that it must orbit the sun.

centered on a line between Earth and the sun. That means it would always be seen as a crescent (Figure 4-10a). But Galileo saw Venus go through a complete set of phases, which proved that it did indeed revolve around the sun (Figure 4-10b). There is no way the Ptolemaic model could produce those phases. This was the strongest evidence that came from Galileo's telescope; but, when controversy erupted, it focused more on the perfection of the sun and moon and the motion of the satellites of Jupiter.

Sidereus Nuncius was very popular and made Galileo famous. He became chief mathematician and philosopher to the Grand Duke of Tuscany in Florence. In 1611, Galileo visited Rome and was treated with great respect. He had long, friendly discussions with the powerful Cardinal Barberini, but he also made enemies. Personally, Galileo was outspoken, forceful, and sometimes tactless. He enjoyed debate, but most of all he enjoyed being right. In lectures, debates, and letters he offended important people who questioned his telescopic discoveries.

By 1616, Galileo was the center of a storm of controversy. Some critics said he was wrong, and others said he was lying. Some refused to look through a telescope lest it mislead them, and others looked and claimed to see nothing (hardly surprising, given the awkwardness of those first telescopes). Pope Paul V decided to end the disruption, so when Galileo visited Rome in 1616 Cardinal Bellarmine interviewed him privately and ordered him to cease debate. There is some controversy today about the nature of Galileo's instructions, but he did not pursue astronomy for some years after the interview. Books relevant to Copernicanism were banned in all Catholic lands, although *De Revolutionibus,* recognized as an important and useful book in astronomy, was only suspended pending revision. Everyone who owned a copy of the book was required to cross out certain statements and add handwritten corrections stating that Earth's motion and the central location of the sun were only theories and not facts.

Dialogo and Trial

In 1621 Pope Paul V died, and his successor, Pope Gregory XV, died in 1623. The next pope was Galileo's acquaintance Cardinal Barberini, who took the name Urban VIII. Galileo rushed to Rome hoping to have the prohibition of 1616 lifted; and, although the new pope did not revoke the orders, he did apparently encourage Galileo. Soon after returning home, Galileo began to write a book on Copernicanism, finally completing it at the end of 1629. After some delay, the book was approved by both the local censor in Florence and the head censor of the Vatican in Rome. It was printed in February 1632.

Called *Dialogo Sopra i Due Massimi Systemi del Mondo (Dialogue Concerning the Two Chief World Systems),* it confronts the ancient astronomy of Aristotle and Ptolemy with the Copernican model and with telescopic observations as evidence. Galileo wrote the book in the form of a debate among three friends. Salviati, a swift-tongued defender of Copernicus, dominates the book; Sagredo is intelligent but largely uninformed. Simplicio, the dismal defender of Ptolemy, makes all the old arguments and sometimes doesn't seem very bright.

The publication of *Dialogo* created an uproar, and it was sold out by August 1632, when the Inquisition ordered sales stopped. The book was a clear defense of Copernicus, and, probably unintentionally, Galileo exposed the pope's authority to ridicule. Urban VIII was fond of arguing that, as God was omnipotent,

He could construct the universe in any form while making it appear to humans to have a different form, and thus its true nature could not be deduced by mere observation. Galileo placed the pope's argument in the mouth of Simplicio, and Galileo's enemies showed the passage to the pope as an example of Galileo's disrespect. The pope thereupon ordered Galileo to face the Inquisition.

Galileo was interrogated by the Inquisition four times and was threatened with torture. He must have thought often of Giordano Bruno, a philosopher, poet, and Dominican monk, who was tried, condemned, and burned at the stake in Rome in 1600. One of Bruno's offenses had been Copernicanism. However, Galileo's trial did not center on his support for Copernicanism. *Dialogo* had been approved by two censors. Rather, the trial centered on the instructions given Galileo in 1616. From his file in the Vatican, his accusers produced a record of the meeting between Galileo and Cardinal Bellarmine that included the statement that Galileo was "not to hold, teach, or defend in any way" the principles of Copernicus. Some historians believe that this document, which was signed neither by Galileo nor by Bellarmine nor by a legal secretary, was a forgery. Others suspect it may be a draft that was never used. It is quite possible that Galileo's actual instructions were much less restrictive; but, in any case, Bellarmine was dead and could not testify at Galileo's trial.

The Inquisition condemned Galileo not for heresy but for disobeying the orders given him in 1616. On June 22, 1633, at the age of 70, kneeling before the Inquisition, Galileo read a recantation admitting his errors. Tradition has it that as he rose he whispered *"E pur si muove"* ("Still it moves"), referring to Earth.

Although he was sentenced to life imprisonment, he was, perhaps through the intervention of the pope, confined at his villa for the next ten years. He died there on January 8, 1642, 99 years after the death of Copernicus.

Galileo was not condemned for heresy, nor was the Inquisition interested when he tried to defend Copernicanism. He was tried and condemned on a charge you might call a technicality. Nevertheless, in his recantation he was forced to abandon all belief in heliocentrism. His trial has been held up as an example of the suppression of free speech and free inquiry and as a famous attempt to deny reality. Some of the world's greatest authors, including Bertolt Brecht, have written about Galileo's trial. That is why Pope John Paul II created a commission in 1979 to re-examine the case against Galileo.

To understand the trial, you must recognize that it was the result of a conflict between two ways of understanding the universe. Since the Middle Ages, biblical scholars had taught that the only path to true understanding was through religious faith. St. Augustine (354–430) wrote *"Credo ut intelligam,"* which can be translated as "Believe in order to understand." Galileo and other scientists of the Renaissance, however, used their own observations as evidence to try to understand nature. When their observations contradicted Scripture, they assumed that their observations represented reality. Galileo paraphrased Cardinal Baronius in saying, "The Bible tells us how to go to heaven, not how the heavens go." The trial of Galileo was not really about the place of Earth in the universe. It was not about Copernicanism. It wasn't even about the instructions Galileo received in 1616. It was, in a larger sense, about the birth of modern science as a rational way to understand the universe (■ Figure 4-11).

The commission appointed by John Paul II in 1979, reporting its conclusions in October 1992, said of Galileo's inquisitors, "This subjective error of judgment, so clear to us today, led them to a disciplinary measure from which Galileo 'had much to suffer.'" Galileo was not found innocent in 1992 so much as the Inquisition was forgiven for having charged him in the first place.

Fair Use

■ **Figure 4-11**

Although he did not invent it, Galileo will always be remembered along with the telescope because it was the source of the evidence he used to try to understand the universe. By depending on direct observation of reality instead of the first principles of philosophy and theology, Galileo led the way to the invention of modern science as a way to know about the natural world.

SCIENTIFIC ARGUMENT

How were Galileo's observations of the moons of Jupiter evidence against the Ptolemaic model?

Scientific arguments are based on evidence, and reasoning from evidence was Galileo's fundamental way of knowing about the heavens. Galileo presented his arguments in the form of evidence that tested the Ptolemaic and Copernican theories, and the moons of Jupiter were key evidence. Ptolemaic astronomers argued that Earth could not move or it would lose its moon, but even in the Ptolemaic universe Jupiter moved, and the telescope showed that it had moons and kept them. Evidently, Earth could move and not leave its moon behind. Furthermore, moons circling Jupiter did not fit the classical belief that all motion was centered on Earth. Obviously there could be other centers of motion. Finally, the orbital periods of the moons are related to their distance from Jupiter, just as in the Copernican system the orbital periods of the planets are related to their distance from the sun. This similarity suggested that the sun rules its harmonious family of planets just as Jupiter rules its harmonious family of moons.

Of all of Galileo's telescopic observations, the moons of Jupiter caused the most debate, but the craters on the moon and the phases of Venus were also critical evidence. Build an argument to discuss that evidence. **How did craters on the moon and the phases of Venus argue against the Ptolemaic model?**

4-5 Isaac Newton and Orbital Motion

The birth of modern astronomy and of modern science date from the 99 years between the deaths of Copernicus and Galileo. The Renaissance is commonly taken to be the period between 1300 and 1600, and that places the 99 years of this story at the culmination of the reawakening of learning in all fields (■ Figure 4-12). Not only did the world adopt a new model of the universe, but it also adopted a new way of understanding humanity's place in nature.

The problem of the place of Earth was resolved by the Copernican Revolution, but the problem of planetary motion was only partly solved by Kepler's laws. For the last 10 years of his life, Galileo studied the nature of motion, especially the accelerated motion of falling bodies. Although he made some important progress, he was not able to relate his discoveries about motion on Earth to that in the heavens. That final step fell to Isaac Newton.

Isaac Newton

Isaac Newton was born in Woolsthorpe, England, on December 25, 1642, and on January 4, 1643. This was not a biological anomaly but a calendrical quirk. Most of Europe, following the lead of the Catholic countries, had adopted the Gregorian calendar, but Protestant England continued to use the Julian calendar. So December 25 in England was January 4 in Europe. If you use the English date, then Newton was born in the same year that Galileo Galilei died.

Newton was a quiet child from a farming family, but his work at school was so impressive that his uncle financed his education at Trinity College, where he studied mathematics and physics. In 1665, plague swept through England, and the colleges were closed. During 1665 and 1666, Newton spent his time at home in Woolsthorpe, thinking and studying. It was during these years that he made most of his discoveries in optics, mechanics, and mathematics. Among other things, he studied optics, developed three laws of motion, divined the nature of gravity, and invented calculus. The publication of his work in his book *Principia* in 1687 placed science on a firm analytical base (■ Figure 4-13).

It is beyond the scope of this book to analyze all of Newton's work, but his laws of motion and gravity shaped the future of astronomy. From his study of the work of Galileo, Kepler, and others, Newton extracted three laws that relate the motion of a body to the forces acting on it (■ Table 4-2). These laws made it possible to predict exactly how a body would move if the forces were known (**How Do We Know? 4-3**).

When Newton thought carefully about motion, he realized that some force must pull the moon toward Earth's center. As described in his first law of motion, an object in motion must continue in motion in a straight line unless it is acted on by a force. If there were no force altering the moon's motion, it would continue moving in a straight line and leave Earth forever. It can circle Earth only if Earth attracts it. Newton's insight was to recognize that the force that holds the moon in its orbit is the same as the force that makes apples fall from trees—gravity.

Newtonian gravitation is sometimes called universal mutual gravitation. Newton's third law points out that forces occur in pairs, so if one body attracts another, the second body must also attract the first. Thus gravitation must be mutual. Furthermore, gravity must be universal. That is, all masses must attract all other masses in the universe. The force between two bodies depends on the masses of the bodies and the distance between them.

The **mass** of an object is a measure of the amount of matter in the object, usually expressed in kilograms. Mass is not the same as weight. An object's weight is the force that Earth's gravity exerts on the object. An object in space far from Earth would have no weight, but it would contain the same amount of matter and would thus have the same mass that it had on Earth.

Newton realized that, in addition to mass, the distance between two objects affects the gravitational attraction between them. He recognized that the force of gravity decreases as the square of the distance between the objects increases. Specifically, if the distance from, say, Earth to the moon were doubled, the gravitational force between them would decrease by a factor of 2^2, which equals 4. If the distance were tripled, the force would decrease by a factor of 3^2, which equals 9. This relationship is known as the **inverse square relation**. (This relation is discussed in Chapter 13, where it is applied to the intensity of light.)

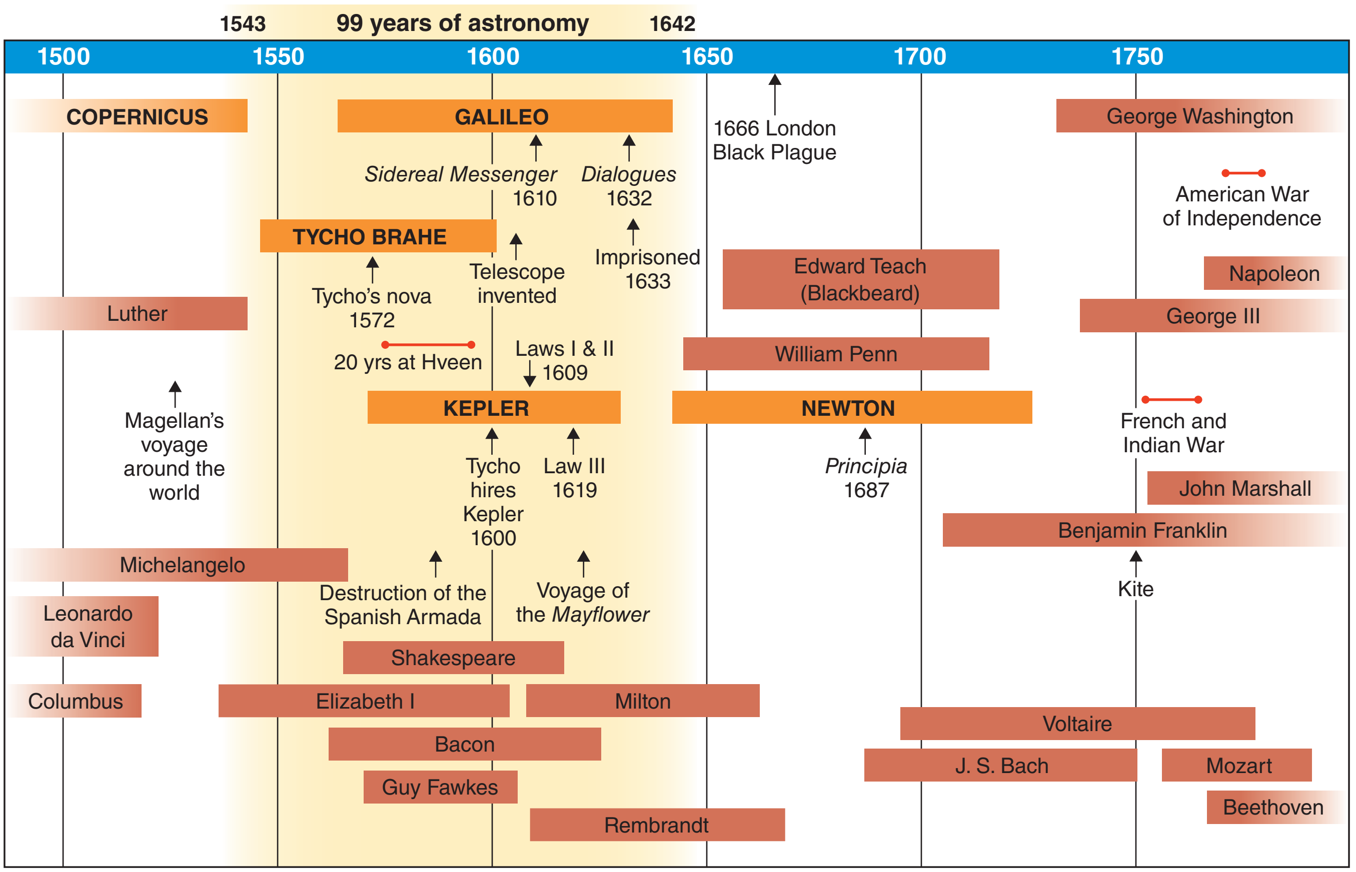


Figure 4-12

The 99 years between the death of Copernicus in 1543 and the birth of Newton in 1642 marked the transition from the ancient astronomy of Aristotle and Ptolemy to the revolutionary hypothesis of Copernicus and, simultaneously, the invention of science as a way to understand nature.

Fair Use

Figure 4-13

Isaac Newton, working from the discoveries of Galileo and Kepler, derived three laws of motion and the principle of mutual gravitation. He and some of his discoveries were honored on this old English 1-pound note. Notice the diagram of orbital motion in the background and the open copy of his book, *Principia,* in his hands.

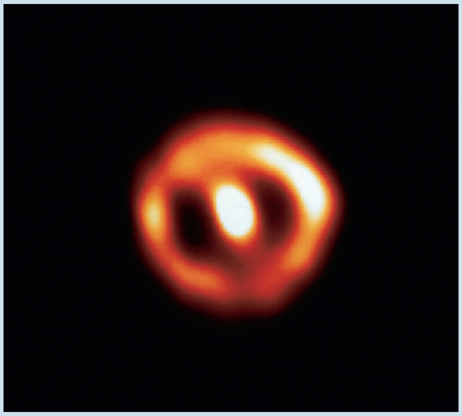
How Do We Know? 4-3

Cause and Effect

Why is the principle of cause and effect so important to scientists? One of the most often used and least often stated principles of science is cause and effect. Modern scientists all believe that events have causes, but ancient philosophers such as Aristotle argued that objects moved because of tendencies. They said that earth and water, and objects made mostly of earth and water, had a natural tendency to move toward the center of the universe. This natural motion had no cause but was inherent in the nature of the objects. Newton's second law of motion *(F = ma)* was the first clear statement of the principle of cause and effect. If an object (of mass *m*) changes its motion (*a* in the equation), then it must be acted on by a force (*F* in the equation). Any effect *(a)* must be the result of a cause *(F)*.

The principle of cause and effect goes far beyond motion. It gives scientists confidence that every effect has a cause. The struggle against disease is an example. Cholera is a horrible disease that can kill its victims in hours. Long ago it was probably blamed on bad magic or the will of the gods, and only two centuries ago it was blamed on "bad air." When an epidemic of cholera struck England in 1854, Dr. John Snow carefully mapped cases on a map of London showing that the victims had drunk water from a small number of wells. Water from those wells caused cholera, and it was soon found that the wells were contaminated by sewage. In 1876, the German Dr. Robert Koch traced cholera to an even more specific cause when he identified the microscopic bacillus that causes the disease. Step by step, scientists tracked down the cause of cholera.

If the universe did not depend on cause and effect, then you could never expect to understand how nature works. Newton's second law of motion was arguably the first clear statement that the behavior of the universe depends rationally on causes.


ESA/STScI and NASA

Cause and effect: Why did this star explode in 1992? There must have been a cause.

Table 4-2 | Newton's Three Laws of Motion

I. A body continues at rest or in uniform motion in a straight line unless acted upon by some force.

II. The change of motion *(a)* of a body of mass *m* is proportional to the force *(F)* acting on it and is in the direction of the force.

$$F = ma$$

III. When one body exerts a force on a second body, the second body exerts an equal and opposite force back on the first body.

With these definitions of mass and the inverse square relation, you can describe Newton's law of gravity in a simple equation:

$$F = -G\frac{Mm}{r^2}$$

Here F is the force of gravity acting between two objects of mass M and m, and r is the distance between their centers. G is the gravitational constant, just a number that depends on the units used for mass, distance, and force. The minus sign reminds you that the force is attractive, tending to make r decrease. To summarize, the force of gravity attracting two objects to each other equals the gravitational constant times the product of their masses divided by the square of the distance between the objects.

Orbital Motion

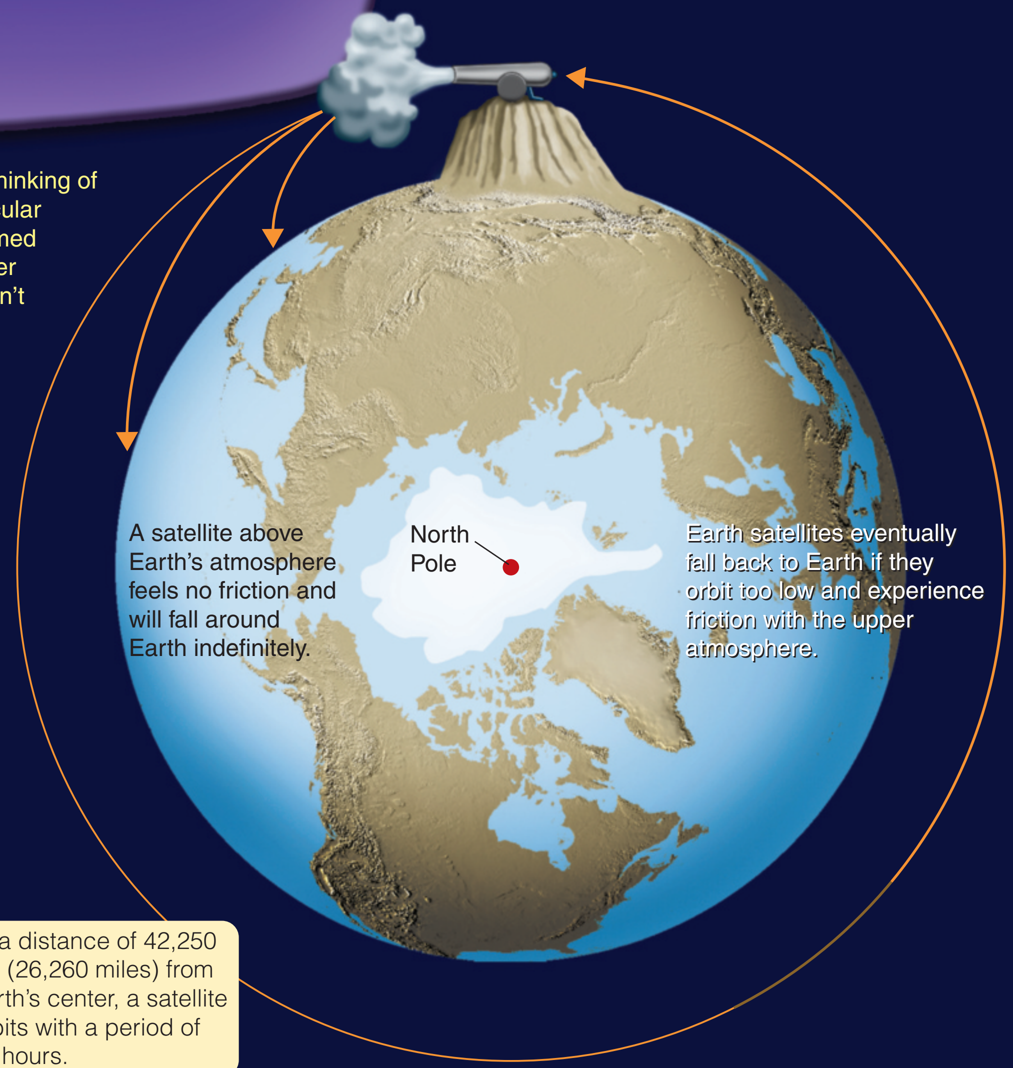
Newton's laws of motion and gravitation make it possible to understand why the moon orbits Earth and how the planets move along their orbits around the sun. You can even discover why Kepler's laws work.

To understand how an object can orbit another object, it helps to describe orbital motion as Newton did—as a form of falling. Study **Orbiting Earth** on pages 66–67 and notice three important ideas and six new terms:

1. An object orbiting Earth is actually falling (being accelerated) toward Earth's center. The object continuously misses Earth because of its orbital velocity. To follow a circular orbit, the object must move with *circular velocity,* and at the right distance from Earth it could be a very useful *geosynchronous satellite.*
2. Notice that objects orbiting each other actually revolve around their *center of mass.*
3. Finally, notice the difference between *closed orbits* and *open orbits.* If you want to leave Earth never to return, you must give your spaceship at least *escape velocity, Ve,* so it will follow an open orbit.

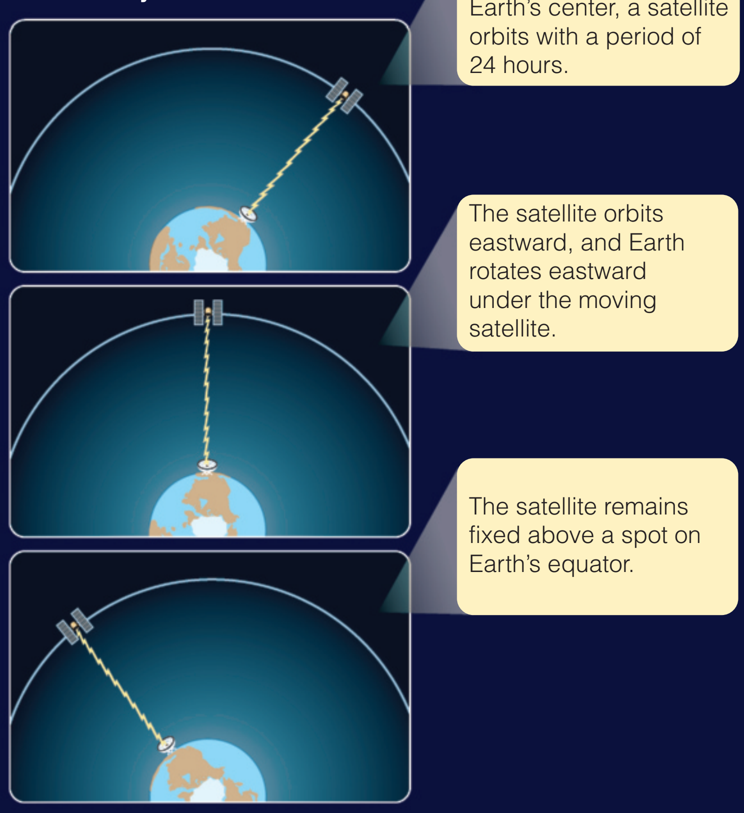
Orbiting Earth

1 You can understand orbital motion by thinking of a cannonball falling around Earth in a circular path. Imagine a cannon on a high mountain aimed horizontally as shown at right. A little gunpowder gives the cannonball a low velocity, and it doesn't travel very far before falling to Earth. More gunpowder gives the cannonball a higher velocity, and it travels farther. With enough gunpowder, the cannonball travels so fast it never strikes the ground. Earth's gravity pulls it toward Earth's center, but Earth's surface curves away from it at the same rate it falls. It is in orbit. The velocity needed to stay in a circular orbit is called the **circular velocity.** Just above Earth's atmosphere, circular velocity is 7790 m/s or about 17,400 miles per hour, and the orbital period is about 90 minutes.

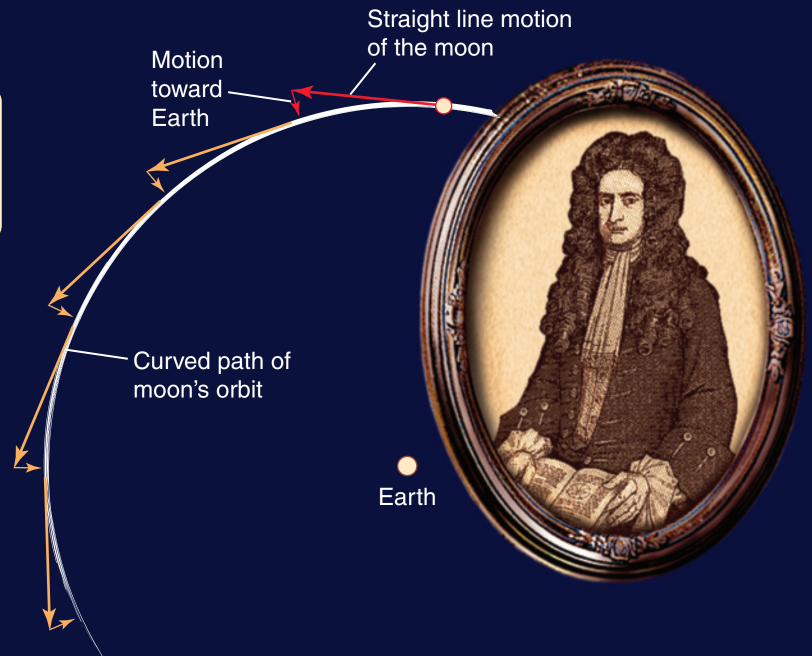


1a A **geosynchronous satellite** orbits eastward with the rotation of Earth and remains above a fixed spot — ideal for communications and weather satellites.

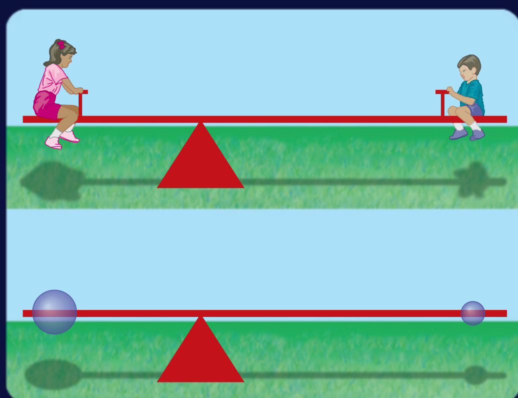
A Geosynchronous Satellite



1b According to Newton's first law of motion, the moon should follow a straight line and leave Earth forever. Because it follows a curve, Newton knew that some force must continuously accelerate it toward Earth — gravity. Each second the moon moves 1020 m (3350 ft) eastward and falls about 1.6 mm (about 1/16 inch) toward Earth. The combination of these motions produces the moon's curved orbit. The moon is falling.

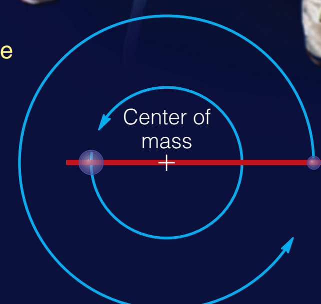


1c Astronauts in orbit around Earth feel weightless, but they are not "beyond Earth's gravity," to use a term from old science fiction movies. Like the moon, the astronauts are accelerated toward Earth by Earth's gravity, but they travel fast enough along their orbits that they continually "miss the Earth." They are literally falling around Earth. Inside or outside a spacecraft, astronauts feel weightless because they and their spacecraft are falling at the same rate. Rather than saying they are weightless, you should more accurately say they are in free fall.

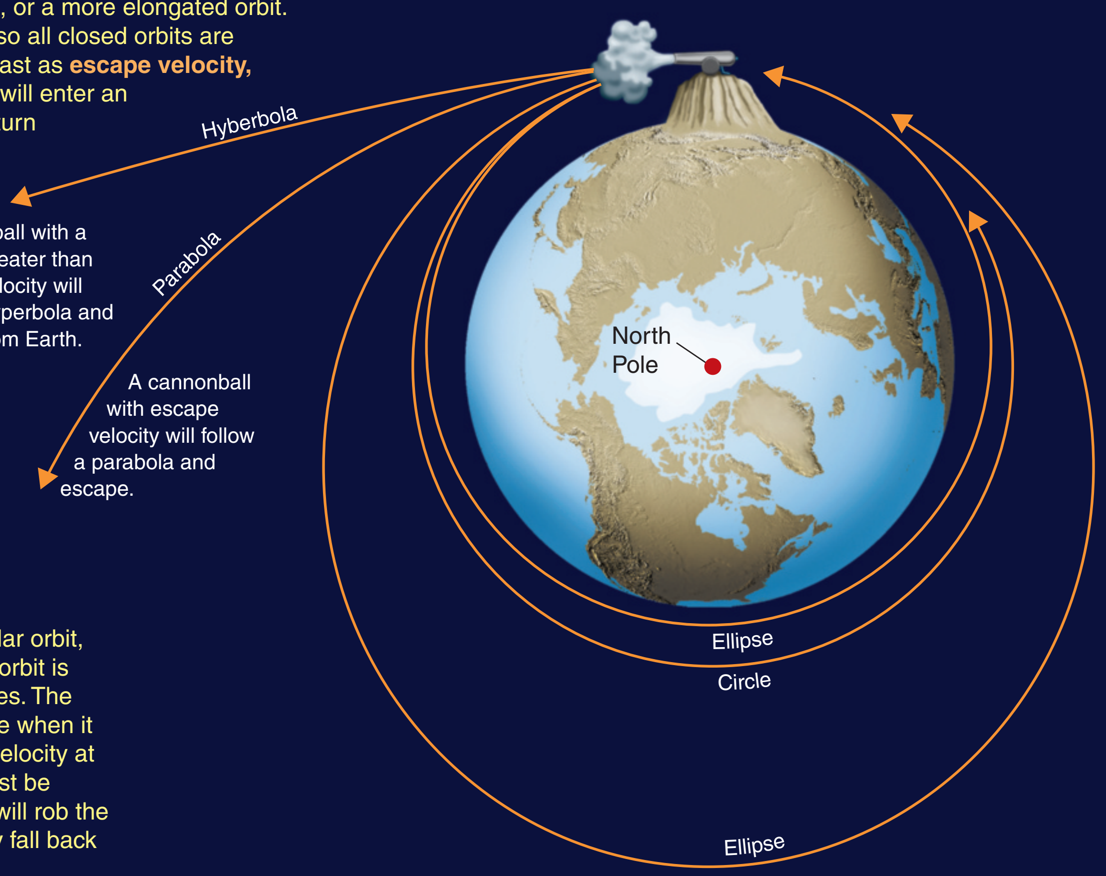


2 To be precise you should not say that an object orbits Earth. Rather the two objects orbit each other. Gravitation is mutual, and if Earth pulls on the moon, the moon pulls on Earth. The two bodies revolve around their common **center of mass,** the balance point of the system.

2a Two bodies of different mass balance at the center of mass, which is located closer to the more massive object. As the two objects orbit each other, they revolve around their common center of mass as shown at right. The center of mass of the Earth–moon system lies only 4708 km (2926 miles) from the center of Earth — inside the Earth. As the moon orbits the center of mass on one side, the Earth swings around the center of mass on the opposite side.



3 **Closed orbits** return the orbiting object to its starting point. The moon and artificial satellites orbit Earth in closed orbits. Below, the cannonball could follow a circular orbit, or a more elongated orbit. (Recall that a circle is also an ellipse, so all closed orbits are elliptical.) If the cannonball travels as fast as **escape velocity,** the velocity needed to leave a body, it will enter an open orbit. An **open orbit** does not return the cannonball to Earth. It will escape.



3a If an object is in a perfectly circular orbit, it's velocity is constant, but if it's orbit is elongated, Kepler's Second Law applies. The object has its lowest velocity at apogee when it is farthest from Earth and its highest velocity at perigee when it is closest. Perigee must be above Earth's atmosphere, or friction will rob the satellite of energy and it will eventually fall back to Earth.

When the captain of a spaceship says, "Put us into a circular orbit," the ship's computers must quickly calculate the velocity needed to achieve a circular orbit. That circular velocity depends only on the mass of the planet and the distance from the center of the planet (**Reasoning with Numbers 4-1**). Once the engines fire and the ship reaches circular velocity, the engines can shut down. The spaceship is then in orbit and will fall around the planet forever so long as it is above the atmosphere where there is no friction. No further effort is needed to maintain orbit, thanks to Newton's laws.

You have probably met a **Common Misconception** if you watch science fiction movies. People in spaceships are usually shown walking around as if they had gravity holding them to the floor. Of course, they should be floating in free fall in their spaceships, unless the rockets are firing, in which case the crew should be strapped into their seats. Authors invent artificial gravity to explain this problem away, but no physicist has ever found a way to generate artificial gravity.

Reasoning with Numbers | 4-1

Circular Velocity

Circular velocity is the velocity a satellite must have to remain in a circular orbit around a larger body. If the mass of the satellite is small compared with the central body, then the circular velocity is given by

$$V_c = \sqrt{\frac{GM}{r}}$$

In this formula, M is the mass of the central body in kilograms, r is the radius of the orbit in meters, and G is the gravitational constant, 6.67×10^{-11} m^3/s^{2}kg. This formula is all you need to calculate how fast an object must travel to stay in a circular orbit.

For example, how fast does the moon travel in its orbit? The mass of Earth is 5.98×10^{24} kg, and the moon orbits 3.84×10^8 m from Earth's center. The moon's velocity is

$$V_c = \sqrt{\frac{6.67 \times 10^{-11} \times 5.98 \times 10^{24}}{3.84 \times 10^8}}$$

$$V_c = \sqrt{\frac{39.9 \times 10^{13}}{3.84 \times 10^8}}$$

$$V_c = \sqrt{1.04 \times 10^6} = 1020 \text{ m/s}$$

This calculation shows that the moon travels 1.02 km along its orbit each second.

Tides

Newton understood that gravity is mutual—Earth attracts the moon, and the moon attracts Earth—and that means the moon's gravity can explain the ocean tides. But Newton also realized that gravitation is universal, and that means there is much more to tides than just Earth's oceans.

Tides are caused by small differences in gravitational forces. For example, Earth's gravity attracts your body downward with a force equal to your weight. The moon is less massive and more distant, so it attracts your body with a force that is a tiny percent of your weight. You don't notice that little force, but Earth's oceans respond dramatically.

The side of Earth that faces the moon is about 4000 miles closer to the moon than is the center of Earth. Consequently, the moon's gravity, tiny though it is at the distance of Earth, is just a bit stronger when it acts on the near side of Earth than on the center. It pulls on the oceans on the near side of Earth a bit more strongly than on Earth's center, and the oceans respond by flowing into a bulge of water on the side of Earth facing the moon. There is also a bulge on the side of Earth that faces away from the moon because the moon pulls more strongly on Earth's center than on the far side. Thus the moon pulls Earth away from the oceans, which flow into a bulge away from the moon as shown at the top of ■ Figure 4-14.

You might wonder: If Earth and the moon accelerate toward each other, why don't they smash together? The answer is that they would collide in about two weeks except that they are orbiting around their common center of mass. The ocean tides are caused by the accelerations Earth and its oceans feel as they move around that center of mass.

A **Common Misconception** holds that the moon's effect on tides means that the moon has an affinity for water—including the water in your body—and, according to some people, that's how the moon makes you behave in weird ways. That's not true. If the moon's gravity only affected water, then there would be only one tidal bulge, the one facing the moon. As you know, the moon's gravity acts on the rock of Earth as well as on water, and that produces the tidal bulge on the far side of Earth. The rocky bulk of Earth responds to these tidal forces, and although you don't notice, Earth flexes, with the mountains and plains rising and falling by a few centimeters in response to the moon's gravitational pull. The moon has no special affinity for water, and, because your body is so much smaller than Earth, any tides the moon raises in your body are immeasurably small. Ocean tides are large because oceans are large.

You can see dramatic evidence of tides if you watch the ocean shore for a few hours. Though Earth rotates on its axis, the

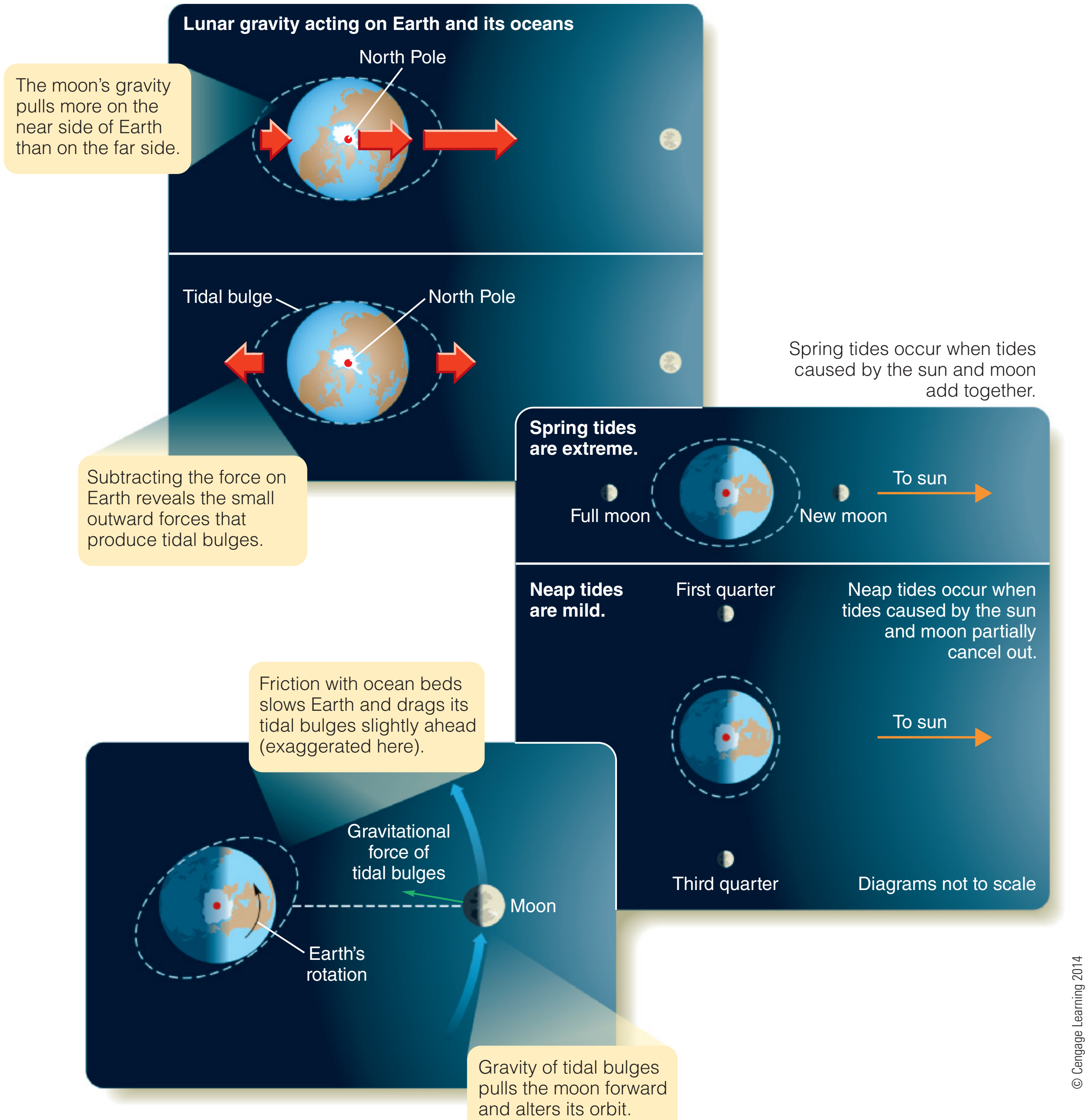


Figure 4-14

Tides are produced by small differences in the gravitational force exerted on different parts of an object. The side of Earth nearest the moon feels a larger force than the side farthest away. Relative to Earth's center, small forces are left over, and they cause the tides. Both the moon and the sun produce tides on Earth; sometimes they add together, and sometimes they partially cancel. Tides can even alter an object's rotation and orbital motion.

tidal bulges remain fixed with respect to the moon. As the turning Earth carries you and your beach into a tidal bulge, the ocean water deepens, and the tide crawls up the sand. The tide does not so much "come in" as you are carried into the tidal bulge. Later, when Earth's rotation carries you out of the bulge, the ocean becomes shallower, and the tide falls. Because there are two bulges on opposite sides of Earth, the tides rise and fall twice a day on an ideal coast.

In reality, the tidal cycle at any given location can be quite complex because of the latitude of the site, shape of the shore,

winds, and so on. Tides in the Bay of Fundy (New Brunswick, Canada), for example, occur twice a day and can exceed 40 feet. In contrast, the northern coast of the Gulf of Mexico has only one tidal cycle a day of roughly 1 foot.

Gravity is universal, so the sun also produces tides on Earth. The sun is roughly 27 million times more massive than the moon, but it lies almost 400 times farther from Earth. Consequently, tides on Earth caused by the sun are less than half as high as those caused by the moon. Twice a month, at new moon and at full moon, the moon and sun produce tidal bulges that add together and produce extreme tidal changes; high tide is very high, and low tide is very low. Such tides are called **spring tides.** Here the word *spring* does not refer to the season of the year but to the rapid welling up of water. At first- and third-quarter moons, the sun and moon pull at right angles to each other, and the sun's tides cancel out some of the moon's tides. These less-extreme tides are called **neap tides**, and they do not rise very high or fall very low. The word *neap* comes from an obscure Old English word, *nep,* that seems to have meant "lacking power to advance." Spring tides and neap tides are illustrated in Figure 4-14.

Galileo tried to understand tides, but it was not until Newton described gravity that astronomers could analyze tidal forces and recognize their surprising effects. For example, the friction of the tidal bulges with the ocean beds slows Earth's rotation and makes the length of a day grow by 0.0023 seconds per century. Fossils of ancient tide markings confirm that only 900 million years ago Earth's day was 18 hours long. Tidal forces can also affect orbital motion. Earth rotates eastward, and friction with the ocean beds drags the tidal bulges slightly eastward out of a direct Earth–moon line. These tidal bulges are massive, and their gravitational field pulls the moon forward in its orbit, as shown at the bottom of Figure 4-14. As a result, the moon's orbit is growing larger by about 3.8 cm a year, an effect that astronomers can measure by bouncing laser beams off reflectors left on the lunar surface by the Apollo astronauts.

Earth's gravitation exerts tidal forces on the moon, and although there are no bodies of water on the moon, friction within the flexing rock has slowed the moon's rotation to the point that it now keeps the same face toward Earth.

Newton's gravitation is much more than just the force that makes apples fall. In later chapters, you will see how tides can pull gas away from stars, rip galaxies apart, and melt the interiors of small moons orbiting near massive planets. Tidal forces produce some of the most surprising and dramatic processes in the universe.

The Newtonian Universe

Newton's insight gave the world a new conception of nature. His laws of motion and gravity were general laws that described the motions of all bodies under the action of external forces. In addition, the laws were productive because they made possible specific calculations that could be tested by observation. For example, Newton's laws of motion can be used to derive Kepler's third law from the law of gravity.

Newton's discoveries remade astronomy into an analytical science in which astronomers could measure the positions and motions of celestial bodies, calculate the gravitational forces acting on them, and predict their future motion (**How Do We Know? 4-4**).

Were you to trace the history of astronomy after Newton, you would find scientists predicting the motion of comets, the gravitational interaction of the planets, the orbits of double stars, and so on. Astronomers built on the discoveries of Newton, just as he had built on the discoveries of Copernicus, Tycho, Kepler, and Galileo. It is the nature of science to build on the discoveries of the past, and Newton was thinking of that when he wrote, "If I have seen farther than other men, it is because I stood upon the shoulders of giants."

How Do We Know? 4-4

Testing a Hypothesis by Prediction

How are the predictions of a hypothesis useful in science? Scientific hypotheses face in two directions. They look back into the past and explain phenomena previously observed. For example, Newton's laws of motion and gravity explained observations of the movements of the planets made over many centuries. But hypotheses also look forward in that they make predictions about what you should find as you explore further. For example, Newton's laws allowed astronomers to calculate the orbits of comets, predict their return, and eventually understand their origin.

Scientific predictions are important in two ways. First, if a prediction of a hypothesis is confirmed, scientists gain confidence that the hypothesis is a true description of nature. But predictions are important for a second reason. They can point the way to unexplored avenues of knowledge.

Predictions have played a key role in particle physics. In the early 1970s, physicists proposed a hypothesis about the fundamental forces and particles in atoms called the Standard Model. This hypothesis explained what scientists had already observed in experiments, but it also predicted the existence of particles that hadn't yet been observed. To test the hypothesis, scientists focused their efforts on building more and more powerful particle accelerators in the hopes of detecting the predicted particles.

A number of these particles have since been discovered, and they do match the characteristics predicted by the Standard Model, further confirming the hypothesis. One predicted particle, the Higgs boson, has not yet been found, as of this writing, but an even larger accelerator may allow its detection. Will the Higgs boson be found, or will physicists have to come up with a better hypothesis? This is just one of many cliff-hangers in modern science.

You learned in an earlier chapter that a hypothesis that has passed many tests and made successful predictions can "graduate" to being considered a theory. As you read about any scientific hypothesis or theory, think about both what it can explain and what it can predict.

Courtesy Brookhaven National Laboratory

Physicists build huge accelerators to search for the subatomic particles predicted by their theories.

What Are We? Participants

The scientific revolution began when Copernicus made humanity part of the universe. Before Copernicus, people thought of Earth as a special place different from any of the objects in the sky; but, in trying to explain the motions in the sky, Copernicus made Earth one of the planets. Galileo and those who brought him to trial understood the significance of making Earth just a planet. It made humanity part of nature, part of the universe.

Kepler showed that the planets move, not at the whim of ancient gods, but according to simple rules, and Newton found simple rules that account for the fall of an apple, orbital motion, and the ocean tides. We are not in a special place ruled by mysterious tendencies. Earth, the sun, and all of humanity are part of a universe whose motions can be described by a few fundamental laws. If simple laws describe the motions of the planets, then the universe is not ruled by mysterious influences as in astrology or by the whim of the gods atop Mount Olympus. And if the universe can be described by simple rules, then the principle of cause and effect applies and the universe is open to scientific study.

Before Copernicus, people felt they were special because they thought they were at the center of the universe and separate from the heavens. Copernicus, Kepler, and Newton showed that we are not at the center but are part of an elegant and complex universe. Astronomy tells us that we are special because we can study the universe and eventually understand what we are. But it also tells us that we are not just observers; we are participants.

Study and Review

Summary

- Ancient philosophers accepted as a first principle that the heavens were perfect, so philosophers such as Plato argued that, because the sphere was the only perfect geometrical form and carried a point on its surface around in a circle, the heavens must move in **uniform circular motion (p. 48).**
- They also accepted that Earth was the unmoving center of all motion in the middle of a **geocentric universe (p. 48).** These principles became part of the teachings of the great philosopher Aristotle, who argued that the sun, moon, and stars were carried around Earth on rotating crystalline spheres.
- The lack of any **parallax (p. 48)** in the positions of the stars gave astronomers confidence that Earth could not be moving.
- About 140 CE, Ptolemy gave mathematical form to Aristotle's model in the *Almagest*. Ptolemy preserved the principles of geocentrism and uniform circular motion, but he added **epicycles (p. 49), deferents (p. 49),** and **equants (p. 49)** to better predict the motions of the planets. To account for **retrograde motion (p. 48),** his epicycles had to be quite large. Even so, his model was not very accurate in predicting the positions of the planets.
- The problem of the place of Earth was solved when Copernicus devised a model that was a **heliocentric universe (p. 50).** He preserved the principle of uniform circular motion, but he put the sun at the center and argued that Earth rotates on its axis and circles the sun once a year. His theory was controversial in part because it contradicted Church teaching.
- Copernicus published his theory in his book *De Revolutionibus* in 1543, the same year he died.
- In Copernicus's model, retrograde motion was explained without epicycles, but because he kept uniform circular motion, he had to include small epicycles, and his model did not predict the motions of the planets well.
- One reason the Copernican model won gradual acceptance was that it was more elegant. Venus and Mercury were treated the same as all the other planets, and the velocity of each planet was related to its distance from the sun. The shift from the geocentric **paradigm (p. 53)** to the heliocentric paradigm is an example of a scientific revolution.
- The problem of planetary motion was finally solved though the work of two astronomers, Tycho Brahe and Johannes Kepler.
- Tycho developed his own model in which the sun and moon circled Earth and the planets circled the sun. His great contribution was to compile detailed observations over a period of 20 years, observations that were later used by Kepler.
- Johannes Kepler inherited Tycho's books of observations in 1601 and used them to devise three laws of planetary motion. The first law says that the planets follow **ellipses (p. 56)** with the sun at one focus. According to the second law, planets move faster when nearer the sun and slower when farther away. The third law says that a planet's orbital period squared is proportional to the **semimajor axis (p. 57)** of its orbit cubed.
- The **eccentricity (p. 57)** of an ellipse equals zero for a circle and grows closer and closer to one as the ellipse becomes more and more elongated.
- A **hypothesis (p. 58)** is a statement about nature that needs further testing, but a **theory (p. 58)** is usually a description of nature that has been tested. Some theories are very well understood and widely accepted. A **natural law (p. 58)** is a fundamental principle in which scientists have great confidence.
- Kepler's final book, *The Rudolphine Tables* (1627), combined heliocentrism with elliptical orbits and predicted the positions of the planets well.
- Galileo used the newly invented telescope to observe the heavens, and he recognized the significance of what he saw there. His discoveries of the phases of Venus, the **Galilean moons (p. 60)** of Jupiter, the mountains of the moon, and other phenomena helped undermine the Ptolemaic universe.
- Galileo based his analysis on observational evidence rather than on first principles or on scripture. In 1633, he was condemned before the Inquisition for refusing to halt his defense of Copernicanism.
- Newton used the work of Kepler and Galileo to discover three laws of motion and the law of gravity. These laws made it possible to understand such phenomena as orbital motion and the tides.
- Newton showed that gravity was mutual and universal. It depends on the **mass (p. 63)** of the bodies and the distance between them according to the **inverse square relation (p. 63).**
- Newton used the image of a cannon on a mountaintop to explain that an object in orbit is falling toward Earth's center and simultaneously moving fast enough to continually miss hitting Earth's surface. To maintain a circular orbit, the object must have **circular velocity (p. 66).** All elliptical orbits, including circles, are **closed orbits (p. 67),** but if the object's velocity equals or exceeds **escape velocity (p. 67),** it will follow an **open orbit (p. 67)** and never return.
- **Geosynchronous satellites (p. 66)** orbit far enough from Earth that their orbital period is 24 hours, and they remain above a single spot on Earth as Earth turns.
- Two objects that orbit each other actually orbit their common **center of mass (p. 67).**
- Newton's laws gave scientists a unified way to think about nature—cause and effect. Every effect has a cause, and science is the search for those causes.
- Newton's laws also explain that tides are caused by small differences in the gravity acting on different parts of a body. Ocean tides occur because the moon's gravity pulls more strongly on the near side of Earth than on the center. A tidal bulge occurs on the far side of Earth because the moon's gravity is slightly weaker there than on the center of Earth.
- Tides produced by the moon combine with tides produced by the sun to cause extreme tides, called **spring tides (p. 70),** at new and full moons. The moon and sun work against each other to produce less extreme tides, called **neap tides (p. 70),** at quarter moons.
- Friction from tides can slow the rotation of a rotating world, and the gravitational pull of tidal bulges can make orbits change slowly.
- The 99 years from the death of Copernicus to the birth of Newton marked the beginning of modern science. From that time on, science depended on evidence to test theories and relied on the analytic methods first demonstrated by Kepler and Newton.

Review Questions

1. Why did Greek astronomers conclude that the heavens were made up of perfect crystalline spheres moving at constant speeds?
2. Why did classical astronomers conclude that Earth had to be motionless?
3. How did the Ptolemaic model explain retrograde motion?
4. In what ways were the models of Ptolemy and Copernicus similar?
5. Why did the Copernican hypothesis win gradual acceptance?
6. Why is it difficult for scientists to replace an old paradigm with a new paradigm?
7. Why did Tycho Brahe expect the new star of 1572 to show parallax? Why was the lack of parallax evidence against the Ptolemaic model?
8. How was Tycho's model of the universe similar to the Ptolemaic model? How did it resemble the Copernican model?
9. Explain how Kepler's laws contradict uniform circular motion.
10. What is the difference between a hypothesis, a theory, and a law?
11. How did *The Alfonsine Tables, The Prutenic Tables,* and *The Rudolphine Tables* differ?
12. Review Galileo's telescopic discoveries and explain why they supported the Copernican model and contradicted the Ptolemaic model.
13. Galileo was condemned by the Inquisition, but Kepler, also a Copernican, was not. Why not?
14. How do Newton's laws lead you to conclude that gravitation has to be universal?
15. Explain why you might describe the orbital motion of the moon with the statement, "The moon is falling."
16. **How Do We Know?** How does a paradigm affect the questions you ask and the answers you find acceptable?
17. **How Do We Know?** How would you respond to someone who said, "Oh, that's only a theory."
18. **How Do We Know?** Why would science be impossible if the principle of cause and effect didn't always apply?
19. **How Do We Know?** Why is it important that a theory make testable predictions?

Discussion Questions

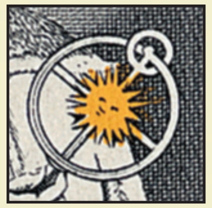
1. Science historian Thomas Kuhn has said that *De Revolutionibus* was a revolution-making book but not a revolutionary book. How was it classical and conservative?
2. Why might Tycho Brahe have hesitated to hire Kepler? Why do you suppose he finally decided to appoint Kepler his scientific heir?
3. Many historians suspect that Galileo offended Pope Urban VIII by putting the pope's favorite argument into the mouth of Simplico. How is the pope's argument a challenge to the principle of cause and effect?
4. How does the modern controversy over creationism and evolution reflect two ways of knowing about the physical world?

Problems

1. If you lived on Mars, which planets would describe retrograde loops? Which would never be visible as crescent phases?
2. Galileo's telescope showed him that Venus has a large angular diameter (61 arc seconds) when it is a crescent and a small angular diameter (10 arc seconds) when it is nearly full. Use the small-angle formula to find the ratio of its maximum distance to its minimum distance. Is this ratio compatible with the Ptolemaic universe shown on page 49? (*Hint:* See Reasoning with Numbers 3-1.)
3. Galileo's telescopes were not of high quality by modern standards. He was able to see the moons of Jupiter, but he never reported seeing features on Mars. Use the small-angle formula to find the maximum angular diameter of Mars when it is closest to Earth. How does that compare with the maximum diameter of Jupiter?
4. If a planet had an average distance from the sun of 10 AU, what would its orbital period be?
5. If a space probe were sent into an orbit around the sun that brought it as close as 0.5 AU to the sun and as far away as 5.5 AU, what would its orbital period be?
6. Neptune orbits the sun with a period of 164.8 years. What is its average distance from the sun?
7. Venus's average distance from the sun is 0.72 AU and Saturn's is 9.54 AU. Calculate the circular orbital velocity of Venus and Saturn around the sun. (*Hints:* The mass of the sun is 1.99×10^{30} kg. An AU is 1.50×10^{11} m.)
8. The circular velocity of Earth around the sun is about 30 km/s. Are the arrows for Venus and Saturn correct in Figure 4-3? (*Hint:* See Problem 7.)
9. What is the orbital velocity of an Earth satellite 42,250 km from Earth's center? How long does it take to circle its orbit once?

Learning to Look

1. What three astronomical objects are represented here? What are the two rings?
2. Why can the object shown at the right be bolted in place and used 24 hours a day without adjustment?
3. Why is it a little bit misleading to say that this astronaut is weightless?



Fair Use

NASA/JSC

Larry Mulvehill/The Image Works

4. If you study the two drawings in the middle of page 48 and relax your eyes, you can merge the two images and see the scene in 3D. What clues does your brain use to give you 3D vision?
5. Mercury's orbit hardly deviates from a circle, but you can tell that it is not a circle with just a glance at Figure 4-7. What gives it away? (Hint: Consider the location of the sun.)

Great Debates

1. *Did Kepler Kill Brahe?* As the legend goes, Brahe died of a burst bladder when he stayed too long at the dinner table one night. However, not all believe this tale. Tycho Brahe was exhumed once and is to be exhumed again. The first exhumation in 1901 found that Brahe had unusually high amounts of mercury in his body at the time of his death, suggesting a death by poison. If true, who poisoned Tycho? The book *Heavenly Intrigue: Johannes Kepler, Tycho Brahe, and the Murder Behind One of History's Greatest Scientific Discoveries,* by Joshua and Anne-Lee Gilder, suggests Kepler. Formulate an opinion on whether Kepler killed Brahe.
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find three sources about Kepler and/or Brahe that support *and* do not support your position.
   c. Cite your sources.

2. *Did Pythagoras, Plato, and Aristotle Force Women into Subservient Roles?* Pythagoras (569–475 BCE) formed a school of philosophy, mathematics, and natural science; he later married Teano, a student of the school. The school is thought to have had both men and women teachers and students. The school made several advances in the study of mathematics, and regardless of who at the school made the discovery, the school (and hence Pythagoras) is given credit. Pythagoras is most noted for the discovery of the Pythagorean theorem. Teano is given credit for some mathematical discoveries but only after her husband died and after she took over the Pythagorean school. Plato was born a few generations later, and his philosophy has roots in the Pythagorean's philosophies. History tells us that Plato's students were both men and women, and one of Plato's famous students was Aristotle. Aristotle wrote several books, including *The Nicomachean Ethics* and *Politics,* where he considered women inferior to men. Women were no longer educated alongside men, and their proper place was now considered to be in the home. Aristotle had great influence on Western thinkers for the next 2000 years. Formulate an opinion on whether Pythagoras, Plato, and Aristotle forced women into subservient roles that lasted into the 19th century. Are women still in subservient roles in the 21st century?
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find three sources about Pythagoreans, Plato, and/or Aristotle that support *and* do not support your stand.
   c. Cite your sources.

3. *Plato's Perfect Universe?* Plato thought celestial bodies were perfectly smooth, traveled in perfect circles, and were perfectly spherical. Is it possible that any celestial body in the universe may be perfect in any way, thereby lending support to the classical astronomers' beliefs?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

4. *Tidal Locking.* The moon always has the same face (that is, the near side) facing Earth and is thus tidally locked to Earth. Earth's current rotation changes the side of Earth that faces the moon, but someday Earth's rotation will slow down to the point that the same side of Earth may always face the moon. Could the tidal locking of Earth to the moon render Earth inhospitable? Could humans adapt because Earth's tidal locking will take millions of years?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Embed copies of your pictures, and explain them.
   c. Cite your sources.

5. *Bye-Bye Moon.* The moon moves away from Earth a few centimeters each year. Over a few billon years, the moon will have moved away from Earth about one-fourth its current distance. Could the moon's retreat render Earth inhospitable? Could humans adapt because the moon's movement will take millions of years? If so, what would humans look like in the future?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Embed copies of your pictures, and explain them.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Retrograde Motion
- Kepler's First and Second Laws
- Newton's Laws
- Geosynchronous Orbits
- Center of Mass

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Lab 6: Tides and Tidal Forces

People living by the sea have noticed the rising and falling of tides since prehistoric times—not just noticed, but sometimes depended on the predictability of tides for their livelihood. Knowing when it is best to head out to sea or to gather certain species of marine creatures requires knowing when to expect the next high or low tide.

For millennia, people along coastlines have understood that tides must have some connection with the moon. On a coastline that is not too complicated, the tidal rhythm is simple: two high tides and two low tides per day. It is easy to notice that one of those high tides occurs about when the moon is at its highest in the sky, and the two low tides occur when the moon is near the horizon. Also, there is a longer cycle: The tidal amplitude (total difference between high and low tide) during a day varies in a monthly cycle, with large "spring" tides (meaning, "springy" or "lively," not the spring season) happening near full and new moon, and smaller-amplitude neap tides happening near first and third quarter phases ("half moons").

With the publication of his book *Principia* in 1687, Isaac Newton invented the modern science of physics. One of his many astonishing achievements in that book was showing how the force of gravity could explain humanity's observations of tidal behavior. In fact, one of the main reasons Newton's *Principia* was immediately convincing to scientists of the day was that it contained a mathematical explanation of the timing and sizes of tides that allowed precise predictions that could be, and were, checked and found to be true.

Implied in Newton's understanding of gravity and tides is that ocean tides are just a particular manifestation of a broader phenomenon. In general, the gravity of any body will stress and stretch a nearby body because of the difference between the gravitational forces exerted at opposite sides of the body. This is what an astronomer means by "tides"; not the rising and falling of the ocean during the day but the ultimate cause, that the moon's gravity is stretching Earth and the oceans responds more readily and obviously than the land. But the land does respond. Even if you are living thousands of miles from the nearest ocean, the ground on which you are standing and walking flexes upward and downward a few millimeters in a daily cycle of two high and two low tides, measured and verified by sensitive instruments.

Also implied in Newton's more profound understanding of tides is the reason the moon has stopped turning and faces only one side toward Earth (termed "synchronous rotation"), which you learned about in a previous chapter. Earth's gravity raises tides in the rocky body of the moon, and this stretched shape (with an amplitude measured in kilometers) resists being pointed away from the line toward Earth. A spinning moon would thus experience a type of friction as the moon's tidal bulges tried to point toward and away from Earth while being turned away by rotation. The eventual result, on geological time scales, is that the moon (and, as you will discover later, moons of the other planets) stopped turning relative to its parent planet. And, reciprocally, Earth's rotation is slowing down gradually due to lunar tides.

Sections 1 and 2 of Virtual Astronomy Lab 6, "Tides and Tidal Forces," help you make the connection between Newton's law of gravity and one of its consequences, the phenomena of ocean tides. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

5

Light and Telescopes

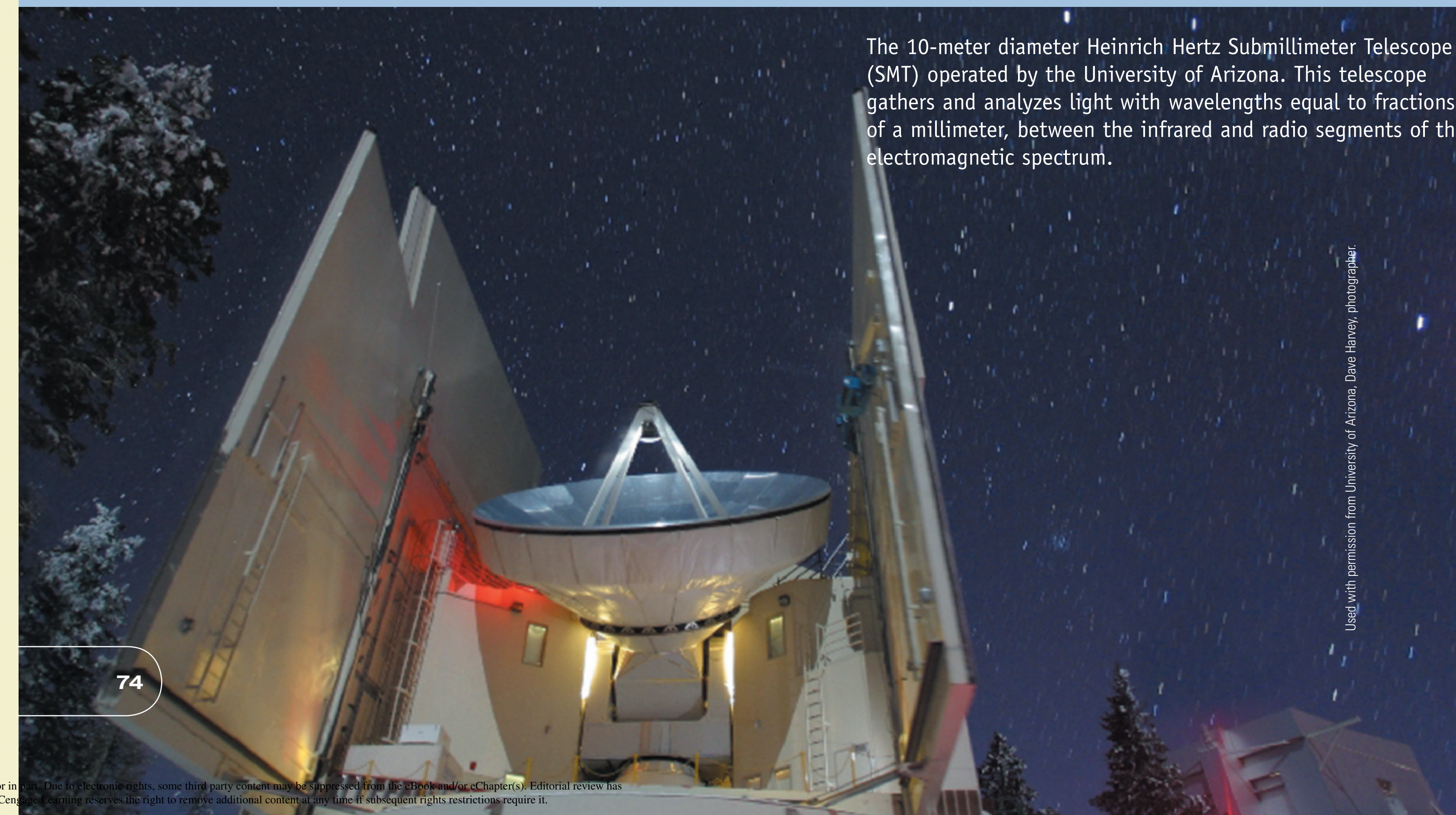
Guidepost

In earlier chapters of this book you looked at the sky the way ancient astronomers did, with the unaided eye. In the previous chapter you got a glimpse through Galileo's telescope that revealed amazing things about the moon, Jupiter, and Venus. Now you can consider the telescopes, instruments, and techniques of modern astronomers.

Telescopes gather and focus light, so you need to study what light is, and how it behaves, on your way to understanding how telescopes work. You will learn about telescopes that capture invisible types of light such as radio waves and X-rays, enabling astronomers to put together a more complete view of celestial objects. This chapter will help you answer these five important questions:

- **What is light?**
- **How do telescopes work?**
- **What are the powers and limitations of telescopes?**
- **What kind of instruments do astronomers use to record and analyze light gathered by telescopes?**
- **Why must some telescopes be located in space?**

The science of astronomy is based on observations. Astronomers cannot visit distant galaxies and far-off worlds, so they have to study them using telescopes. Fifteen chapters remain in your exploration, and every one will present information gained by astronomers using telescopes.



The 10-meter diameter Heinrich Hertz Submillimeter Telescope (SMT) operated by the University of Arizona. This telescope gathers and analyzes light with wavelengths equal to fractions of a millimeter, between the infrared and radio segments of the electromagnetic spectrum.

Used with permission from University of Arizona, Dave Harvey, photographer

The strongest thing that's given us
to see with's
A telescope. Someone in every town
Seems to me owes it to the town to keep one.

ROBERT FROST, "THE STAR-SPLITTER"

Natural light is a treasure that links you to the sky. The quest of astronomers is to study light from the sun, planets, moons, asteroids, comets, stars, nebulae, and galaxies and extract information about their natures. Most celestial objects are very faint sources of light, so large telescopes, like the one shown on the previous page, are built to collect the greatest amount of light possible. Some types of telescopes gather light that is invisible to the human eye, but all telescopes work by the same basic principles. Some telescopes are used on Earth's surface, but others must go high in Earth's atmosphere, or even above the atmosphere into space, to work properly. This chapter focuses on telescopes, instruments and techniques used to analyze light for astronomical research.

Astronomy is more than technology and scientific analysis. It helps us understand what we are. In the quotation that opens this chapter, the poet Robert Frost suggests that someone in every town should own a telescope to help us look upward and outward.

5-1 Radiation: Information from Space

Astronomers no longer spend their time mapping constellations or charting phases of the moon. Modern astronomers analyze light using sophisticated instruments and techniques to investigate the compositions, motions, internal processes, and evolution of celestial objects. To understand how astronomers gain such detailed information about distant objects, you first need to learn about the nature of light.

Light as a Wave and a Particle

When you admire the colors of a rainbow, you are seeing an effect of light acting as a wave (■ Figure 5-1a); when you use a digital camera to take a picture of the same rainbow, the light acts like particles as it hits the camera's detectors (Figure 5-1b). Light has both wavelike and particle-like properties, and how it behaves depends partly on how you treat it.

Light is referred to as **electromagnetic radiation** because it is made up of both electric and magnetic fields. The word "light" is commonly used to refer to electromagnetic radiation that humans can see, but visible light is only one among many types of electromagnetic radiation that include X-rays and radio waves. Some people are wary of the word *radiation,* but that involves a **Common Misconception.** *Radiation* refers to anything that radiates from a source. Dangerous high-energy particles emitted from radioactive atoms are also called radiation, and you have learned to be concerned when you hear that word. But light, like all electromagnetic radiation, spreads outward from its origin, so you can correctly refer to light as a form of radiation.

Electromagnetic radiation travels through space at a speed of 3.00×10^8 m/s (186,000 mi/s), symbolized by the letter *c.* This is commonly referred to as the speed of light, and it is in fact the speed of all types of electromagnetic radiation.

Electromagnetic radiation can act as a wave phenomenon—that is, it is associated with a periodically repeating disturbance, a wave, which carries energy. You are familiar with waves in water: If you disturb a pool of water, waves spread across the surface. Imagine placing a ruler parallel to the travel direction of the wave. The distance between peaks of the wave is called the **wavelength**, usually represented by the Greek lower case letter "lambda" (λ).

Wavelength is related to **frequency**, the number of waves that pass a stationary point in 1 second. Frequency is often represented by the Greek lower case letter "nu" (ν). The relationship between the wavelength, frequency, and speed of a wave can be written in two equivalent ways:

$$\lambda = \frac{c}{\nu} \quad \text{or} \quad \nu = \frac{c}{\lambda}$$

© Hoomar/Shutterstock.com

Roy McMahon

■ **Figure 5-1**

The wavelike properties of light produce a rainbow, while the particle-like properties are involved in the operation of a digital camera.

If your favorite FM station is on the dial at, say, 89.5, that means the station's radio waves have a frequency $\nu = 89.5$ megahertz, in other words, 89.5 million radio wave peaks pass a given point each second. You already know that the radio waves are traveling at the speed of light $c = 3.00 \times 10^8$ m/s. Using the first formula, you can calculate that your favorite radio station is radiating radio waves with a wavelength of $\lambda = 3.35$ m. Note that wavelength and frequency have an inverse relationship: The higher the frequency, the shorter the wavelength.

Sound is another example of a wave: in that case, a periodically repeating pressure disturbance that moves from source to ear. Sound requires a medium, meaning a substance such as air to travel through. On the moon, where there is no air, there can be no sound. In contrast, light is made up of electric and magnetic fields that can travel through empty space. In other words, unlike sound, light does not require a medium and can travel through a perfect vacuum. There is no sound on the moon, but there is plenty of light. A **Common Misconception** is that radio waves are related to sound. Actually, radio waves are a type of light (electromagnetic radiation) that your radio converts into sound so you can listen. Radio communication works just fine between astronauts standing on the airless moon; the spacesuit radios convert the radio signals to sound that is heard in the air inside their helmets.

Although electromagnetic radiation can behave as a wave, it can also behave as a stream of particles. A particle of electromagnetic radiation is called a **photon.** You can think of a photon as a packet of waves. The amount of energy a photon carries is inversely proportional to its wavelength. Notice that there is a relationship between the energy E of a photon, which is a particle property of light, and the wavelength λ, which is a wave property. The inverse proportion means that as λ gets smaller, E gets larger: Shorter-wavelength photons carry more energy, and longer-wavelength photons carry less energy. If you think about it, you will realize that the relationship between wavelength and frequency you learned earlier means there must also be a simple relationship between photon energy and frequency. That is, short wavelength, high frequency, and large photon energy go together, and long wavelength, low frequency, and small photon energy go together.

The Electromagnetic Spectrum

A **spectrum** is an array of electromagnetic radiation displayed in order of wavelength. You are most familiar with the spectrum of visible light that you see in rainbows. The colors of the rainbow differ in wavelength, with red having the longest wavelength and violet the shortest. The visible spectrum is shown in ■ Figure 5-2a.

The average wavelength of visible light is about 0.0005 mm. Roughly 50 light waves would fit end-to-end across the thickness of a sheet of household plastic wrap. It is awkward to describe such short distances in millimeters, so scientists usually give the wavelength of light using **nanometer (nm)** units, equal to one-billionth of a meter (10^{-9} m). Another unit that astronomers commonly use is called the **angstrom (Å),** named after the Swedish astronomer Anders Jonas Ångström. One angstrom is 10^{-10} m, one-tenth of a nanometer. The wavelength of visible light ranges from about 400 to 700 nm (4000 to 7000 Å).

Just as you sense the wavelength of sound as pitch, you sense the wavelength of light as color. Light with wavelengths at the short-wavelength end of the visible spectrum (λ = about 400 nm) looks violet to your eyes, and light with wavelengths at the long-wavelength end (λ = about 700 nm) looks red.

Figure 5-2a shows that the visible spectrum makes up only a small part of the entire electromagnetic spectrum. Beyond the red end of the visible spectrum lies **infrared (IR)** radiation, with wavelengths ranging from 700 nm to about 1 millimeter. Your eyes do not detect infrared, but your skin senses it as heat. A "heat lamp" warms you by giving off infrared radiation. Infrared radiation was discovered as the first known example of "invisible light" in the year 1800 (■ Figure 5-3).

Beyond the infrared part of the electromagnetic spectrum lie **microwaves** and **radio waves**. Microwaves, used for cooking food in a microwave oven, as well as for radar and some long-distance telephone communications, have wavelengths from a few millimeters to a few centimeters. The radio waves used for FM, television, military, government, and cell phone radio transmissions have wavelengths of a few centimeters to a few meters, while AM and other types of radio transmissions have wavelengths of a few hundred meters to a few kilometers.

Now look at the other end of the electromagnetic spectrum in Figure 5-2a and notice that electromagnetic waves shorter than violet are called **ultraviolet (UV)**. Electromagnetic waves that are even shorter are called **X-rays**, and the shortest are **gamma rays**.

Recall the formula for the energy of a photon. Extremely short wavelength, high-frequency photons such as X-rays and gamma rays have high energies and can be dangerous. Even ultraviolet photons have enough energy to harm you. Small amounts of ultraviolet produce a suntan, and larger doses cause sunburn and skin cancers. Contrast this to the lower-energy infrared photons. Individually they have too little energy to affect skin pigment, a fact that explains why you can't get a tan from a heat lamp. Only by concentrating many low-energy photons in a small area, as in a microwave oven, can you transfer significant amounts of energy.

The boundaries between these wavelength ranges are defined only by conventional usage, not by natural divisions. There is no distinction between short-wavelength ultraviolet light and long-wavelength X-rays. Similarly, long-wavelength infrared radiation is indistinguishable from the shortest microwaves.

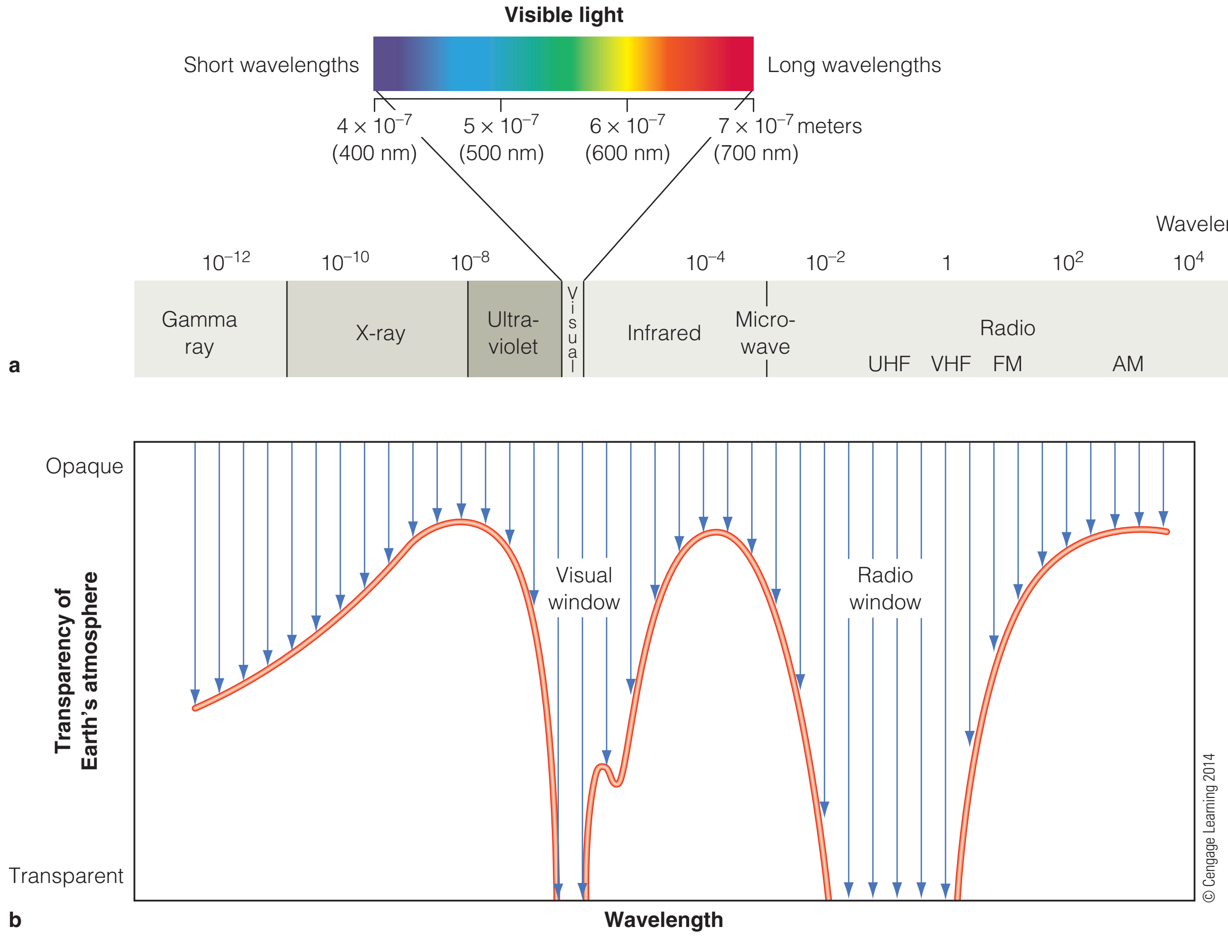


Figure 5-2

(a) The spectrum of visible light, extending from red to violet, is only part of the electromagnetic spectrum. (b) Most forms of light (electromagnetic radiation) are absorbed in Earth's atmosphere. Only light in the visual and radio "windows" can reach Earth's surface easily.

Astronomers capture and study electromagnetic radiation from space because it carries almost the only clues available about the nature of stars, planets, and other celestial objects. Earth's atmosphere is opaque to most electromagnetic radiation, as shown in the graph in Figure 5-2b. Gamma rays and X-rays are absorbed high in Earth's atmosphere, and a layer of ozone (O_3) at altitudes of about 15–30 km absorbs most ultraviolet radiation. Water vapor in the lower atmosphere absorbs most longer-wavelength infrared radiation and microwaves. Only visible light, some shorter-wavelength infrared, and some radio waves reach Earth's surface through wavelength bands called **atmospheric windows**. Obviously, if you wish to study the universe from Earth's surface, you have to "look through" one of those windows.

SCIENTIFIC ARGUMENT

What would you see if your eyes were sensitive only to X-rays? Sometimes critical analysis of an idea is easier if you try to apply it in a totally new situation. In this case, you might at first expect to be able to see through walls, but remember that your eyes detect only light that already exists. There are almost no X-rays flying around at Earth's surface, so if you had X-ray eyes, you would be in the dark and would be unable to see anything. Even when you looked up at the sky, you would see nothing, because Earth's atmosphere is not transparent to X-rays. The comic books seem to show X-rays coming *out* of Superman's eyes, but that's not how vision works. X-ray vision would not help Superman see through a wall unless there was some source of X-rays on the other side.

Now imagine a slightly different situation and modify your argument. **Would you be in the dark if your eyes were sensitive only to radio wavelengths?**

Figure 5-3

Depiction of Sir William Herschel discovering that sunlight contains radiation detectable by thermometers but not by human eyes. He named that invisible light "infrared," meaning, "below red."

5-2 Telescopes

ASTRONOMERS BUILD TELESCOPES to collect light for analysis from distant faint objects. This requires very large telescopes plus careful optical and mechanical engineering work. You can understand these ideas more completely by learning about two types of telescopes and their relative advantages and disadvantages.

Two Ways to Do It: Refracting and Reflecting Telescopes

Light can be focused into an image in one of two ways (■ Figure 5-4). Either: (1) a lens refracts ("bends") the light passing through, or (2) a mirror reflects ("bounces") the light off its surface.

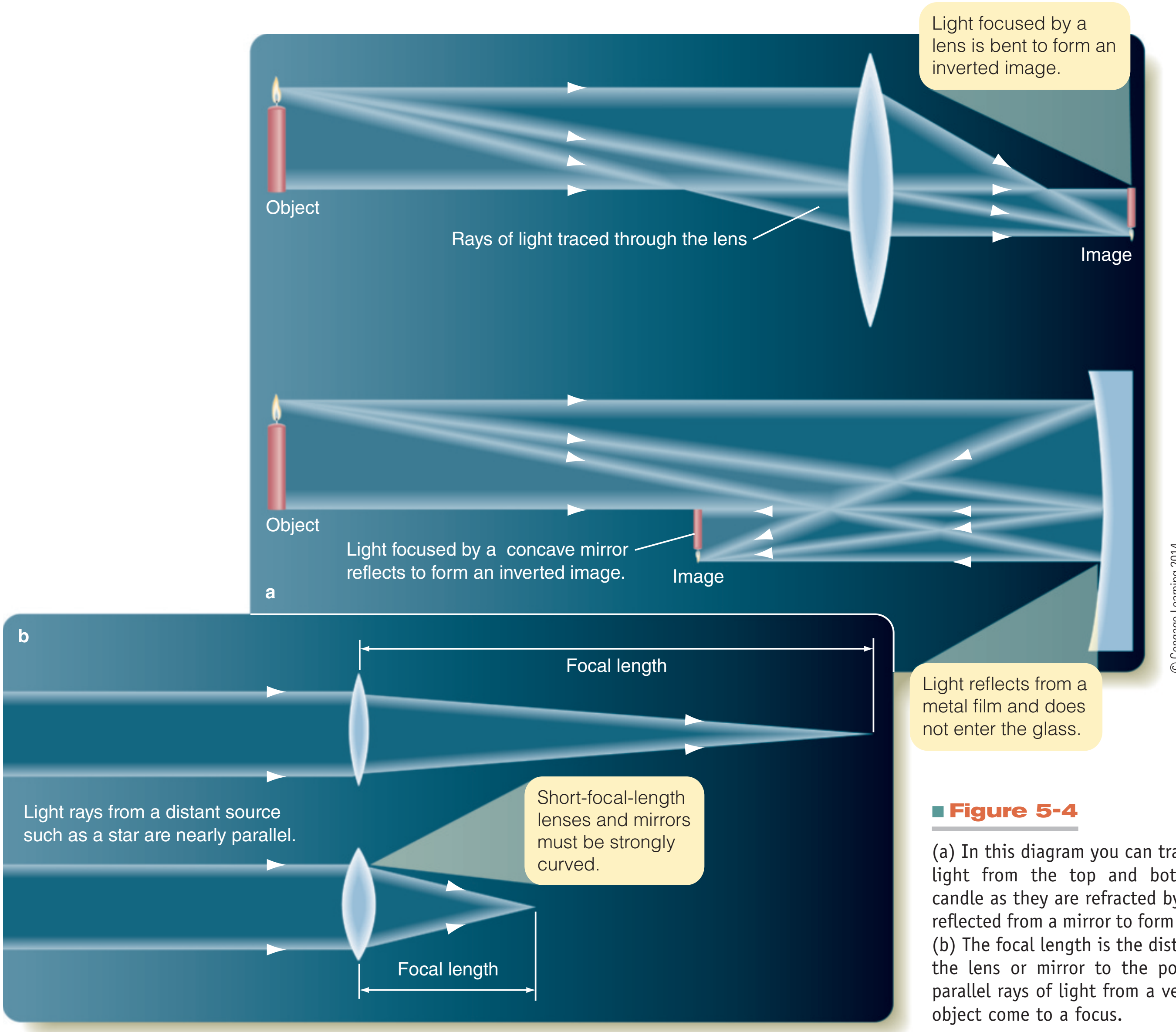
These two ways to manipulate light correspond to two astronomical telescope designs. **Refracting telescopes** use a lens, and **reflecting telescopes** use a mirror, to gather and focus light (■ Figure 5-5). The main lens in a refracting telescope is called the **primary lens**, and the main mirror in a reflecting telescope is called the **primary mirror**. Both kinds of telescopes form an image that is small, inverted, and difficult to observe directly, so a lens called the **eyepiece** normally is used to magnify the image and make it convenient to view. You learned in the previous chapter that Galileo was the first person to systematically record observations of celestial objects using a telescope, beginning almost exactly 400 years ago, in 1610. Galileo's telescope was a refractor. You also learned about the amazing range of Isaac Newton's scientific work; among his many accomplishments was the invention of the reflecting telescope.

The distance from a lens or mirror to the image it forms of a distant light source such as a star is called the **focal length**. Short-focal-length lenses and mirrors must be strongly curved, and long-focal-length lenses and mirrors are less strongly curved. The surfaces of lenses and mirrors must be polished to eliminate irregularities larger than the wavelength of light. Grinding a lens or mirror to the proper shape and necessary smoothness is a delicate, time-consuming, and expensive process. Creating the optics for a large telescope can take months or years; involve huge, precision machinery; and employ several expert optical engineers and scientists.

Refracting telescopes suffer from a serious optical distortion that limits their usefulness. When light is refracted through glass, shorter-wavelength light bends more than longer wavelengths, so, for example, blue light, having shorter wavelengths, comes to a focus closer to the lens than does red light (■ Figure 5-6a). That means if you focus the eyepiece on the blue image, the other colors are out of focus, and you see a colored blur around the image. If you focus instead on the red image, all the colors except red are blurred, and so on. This color separation is called **chromatic aberration**. Telescope designers can grind a telescope lens with two components made of different kinds of glass, and thereby bring two different wavelengths to the same focus (Figure 5-6b). That does improve the image, but these so-called **achromatic lenses** are not totally free of chromatic aberration. Even though two colors have been brought together, the others are still out of focus.

The primary lens of a refracting telescope is much more difficult to manufacture than a mirror of the same size. The interior of the glass must be pure and flawless because the light passes through it. Also, if the lens is achromatic, it must be made of two different kinds of glass requiring four precisely ground surfaces. The largest refracting telescope in the world was completed in 1897 at Yerkes Observatory in Wisconsin. Its achromatic primary lens is 1 m (40 in.) in diameter and weighs half a ton. Refracting telescopes larger than that are prohibitively expensive.

The primary mirrors of reflecting telescopes are much less expensive than lenses because the light reflects off the front surface of the mirror. This means that only the front surface



■ **Figure 5-4**

(a) In this diagram you can trace rays of light from the top and bottom of a candle as they are refracted by a lens or reflected from a mirror to form an image. (b) The focal length is the distance from the lens or mirror to the point where parallel rays of light from a very distant object come to a focus.

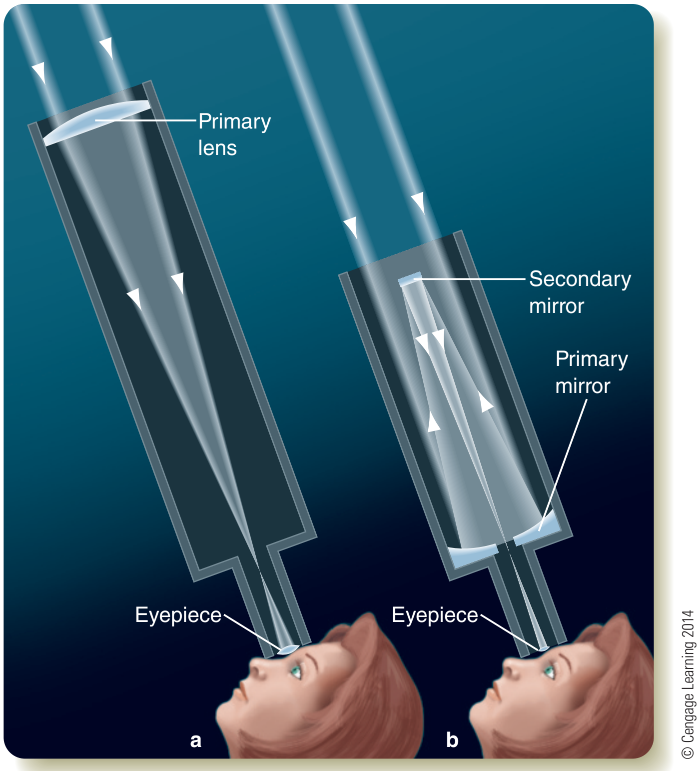
needs to be made with a precise shape. The front surface is coated with a highly reflective surface of aluminum or silver, and the light reflects from that. Consequently, the glass of the mirror does not need to be transparent, and the mirror can be supported across its back surface to reduce bending caused by its own weight. Most important, reflecting telescopes do not suffer from chromatic aberration because the light does not pass through the glass, so reflection does not depend on wavelength. For these reasons, all large astronomical telescopes built since the beginning of the 20th century have been reflecting telescopes.

Telescopes that collect visible light are called **optical telescopes** (■ Figure 5-7a). Like visible light, radio waves from celestial objects can penetrate Earth's atmosphere and reach the ground. Astronomers gather radio waves using **radio telescopes**, such as the one in Figure 5-7b, that resemble giant TV satellite dishes. It is technically very difficult to make a lens that can focus radio waves, so all radio telescopes, including small ones, are reflecting telescopes; the dish is the primary mirror.

Powers and Limitations of Telescopes

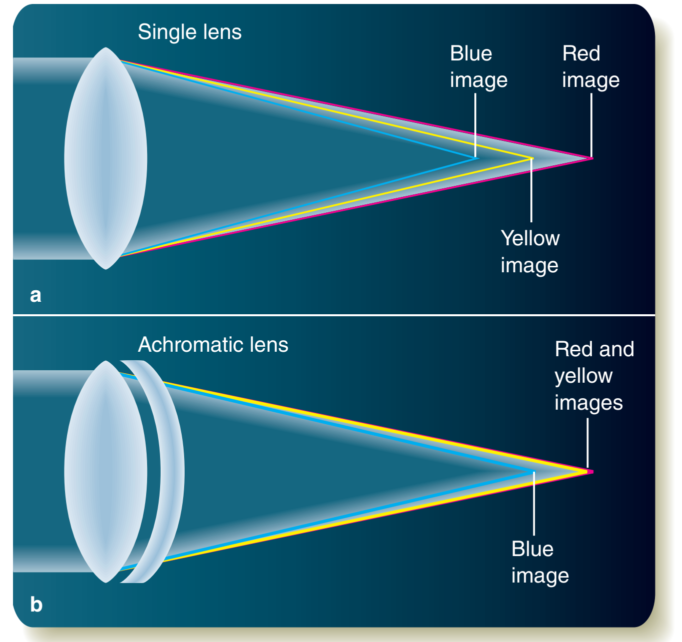
A telescope's capabilities are described in three important ways that are called the three powers of a telescope. The two most important of these powers depends on the diameter of the telescope.

Nearly all of the interesting objects in the sky are faint sources of light, so astronomers need telescopes that can collect large amounts of light to be able to study those objects. **Light-gathering power** refers to the ability of a telescope to collect light. Catching light in a telescope is like catching rain in a bucket—the bigger the



■ **Figure 5-5**

(a) A refracting telescope uses a primary lens to focus starlight into an image that is magnified by another lens called an eyepiece. The primary lens has a long focal length, and the eyepiece has a short focal length. (b) A reflecting telescope uses a primary mirror to focus the light by reflection. In this particular reflector design, called a Cassegrain telescope, a small secondary mirror reflects the starlight back down through a hole in the middle of the primary mirror to the eyepiece lens.



■ **Figure 5-6**

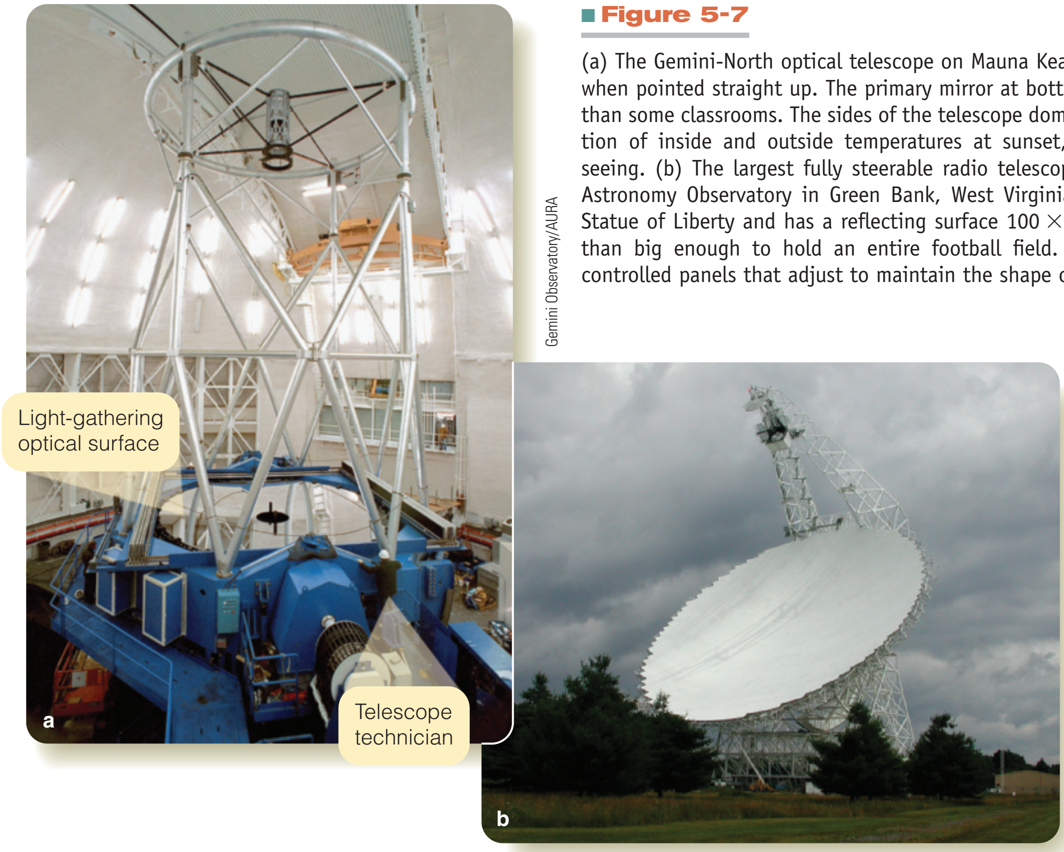
(a) An ordinary lens suffers from chromatic aberration because short wavelengths bend more than long wavelengths. (b) An achromatic lens, with two components made of two different kinds of glass, can bring any two colors to the same focus, but other colors remain slightly out of focus.

bucket, the more rain it catches (■ Figure 5-8). Light-gathering power is proportional to the *area* of the telescope's primary lens or mirror, that is, proportional to the primary's diameter squared. For example, a telescope with a diameter of 2 meters has four times the light-gathering power of a 1-meter telescope. Even a small increase in diameter produces a large increase in light-gathering power and allows astronomers to study much fainter objects. This principle holds not only at visual wavelengths but also for telescopes collecting any kind of light.

The second power of a telescope, called **resolving power**, refers to the ability of the telescope to reveal fine detail. One consequence of the wavelike nature of light is that there is unavoidable blurring called a **diffraction fringe** around every point of light in an image, and you cannot see any detail smaller than that fringe (■ Figure 5-9). Astronomers can't eliminate diffraction fringes, but the size of the diffraction fringes is inversely proportional to the diameter of the telescope. This means that the larger the telescope, the better its resolving power. This principle is also true for telescopes gathering any kind of light. However, the size of the diffraction fringes is also proportional to the wavelength of light being focused. In other words, an infrared or radio telescope has less resolving power than an optical telescope of the same size.

Aside from diffraction, two other factors—optical quality and atmospheric conditions—limit resolving power. A telescope must have high-quality optics to achieve its full potential resolving power. Even a large telescope reveals little detail if its optical surfaces are marred by imperfections. Also, when you look through a telescope, you are looking up through miles of turbulent air in Earth's atmosphere, inevitably making images wiggle and blur to some extent. Astronomers use the term **seeing** to refer to the amount of image blurring due to atmospheric conditions. A related phenomenon is the twinkling of stars. Star twinkles are caused by turbulence in Earth's atmosphere, and a star near the horizon, where you look through more air, will twinkle and blur more than a star overhead.

On a night when the atmosphere is unsteady, images are badly blurred and the seeing is said to be "bad" (■ Figure 5-10). Generally, even under relatively good seeing conditions, the detail visible through a large telescope is limited not by its diffraction fringes but by the turbulence of the air through which the

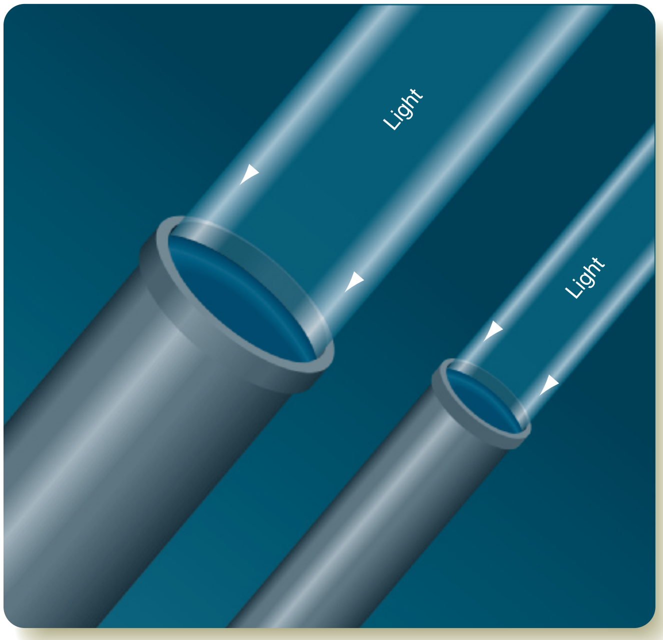


Gemini Observatory/AURA

NRAO/AUI

Figure 5-7

(a) The Gemini-North optical telescope on Mauna Kea in Hawai'i stands over 19 m (60 ft) high when pointed straight up. The primary mirror at bottom is 8.1 m (26.5 ft) in diameter—larger than some classrooms. The sides of the telescope dome can be opened, allowing quick equalization of inside and outside temperatures at sunset, reducing air turbulence and improving seeing. (b) The largest fully steerable radio telescope in the world is at the National Radio Astronomy Observatory in Green Bank, West Virginia. The telescope stands higher than the Statue of Liberty and has a reflecting surface 100 × 110 m (330 × 360 ft) in diameter, more than big enough to hold an entire football field. Its surface consists of 2004 computer-controlled panels that adjust to maintain the shape of the reflecting surface.



© Cengage Learning 2014

Figure 5-8

Gathering light is like catching rain in a bucket. A large-diameter telescope gathers more light and produces a brighter image than a smaller telescope of the same focal length.

telescope must look. An optical telescope performs better on a high mountaintop where the air is thin and steady, but even there Earth's atmosphere spreads star images at visual wavelengths into blobs about 0.5 to 1.0 arc seconds in diameter. Radio telescopes are also affected by atmospheric seeing, but less than optical telescopes, so they do not benefit much in this respect by being located on mountains. You will learn later in this chapter about special techniques that improve seeing from ground-based telescopes, and also about telescopes that orbit above Earth's atmosphere and are not limited by seeing.

Seeing and diffraction limit the amount of information in an image, and that limits the precision of any measurement that can be made using that image. All measurements have some built-in uncertainty (**How Do We Know? 5-1**), and scientists must learn to work within those limitations. Have you ever tried to magnify a newspaper photo to distinguish some detail? Newspaper photos are composed of tiny dots of ink, and no detail smaller than a single dot will be visible no matter how much you magnify the photo. In an astronomical image, the resolution is limited by seeing, or diffraction, or both. You can't see any detail in the image that is smaller than the telescope's resolution. That's why stars look like fuzzy points of light no matter how big the telescope used to view them is.

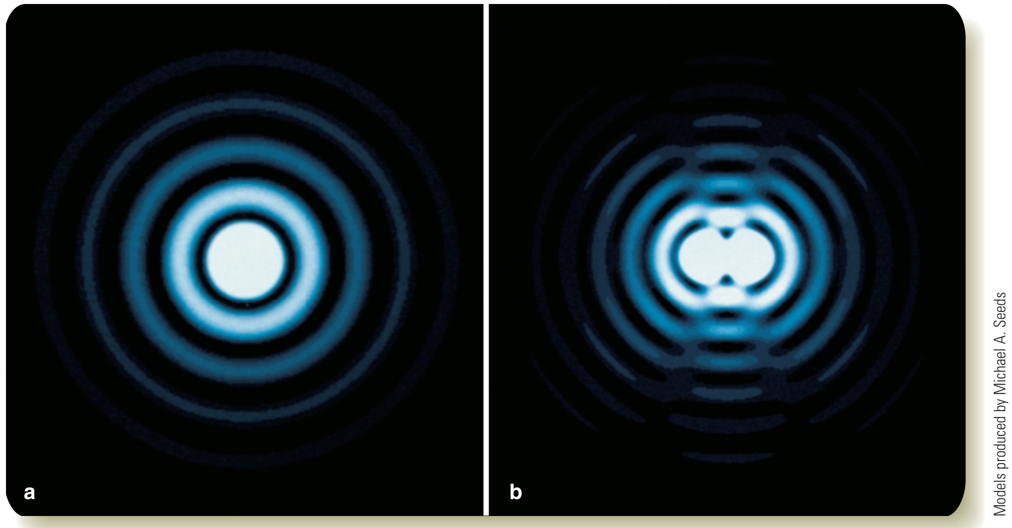


Figure 5-9

(a) Stars are so far away that their images are points, but the wavelike characteristic of light causes each star image to be surrounded with diffraction fringes, much magnified in these computer models. (b) Two stars close to each other have overlapping diffraction fringes and become impossible to detect separately.

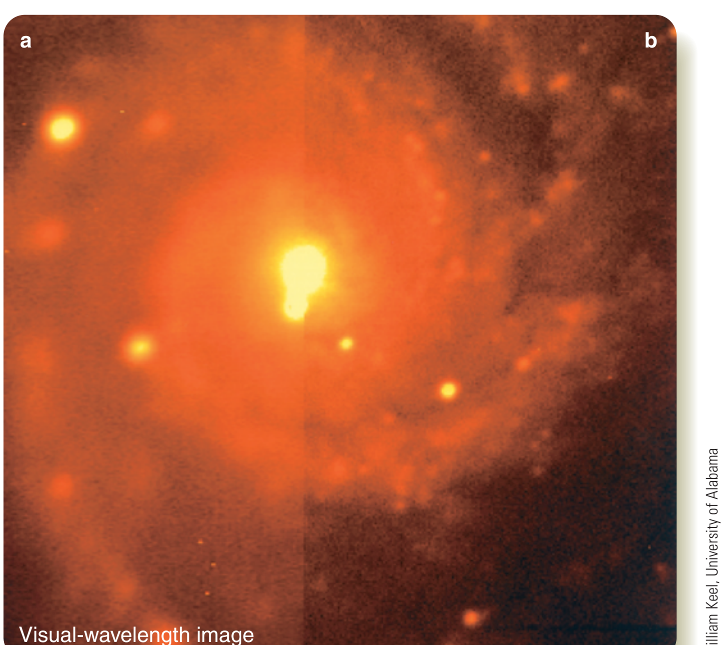


Figure 5-10

(a) The left half of this photograph of a galaxy is from an image recorded on a night of poor seeing. Small details are blurred. (b) The right half of the photo is from an image recorded on a night when Earth's atmosphere above the telescope was steady and the seeing was better. Much more detail is visible under good seeing conditions.

It is a **Common Misconception** that the purpose of an astronomical telescope is to magnify images. In fact, the **magnifying power** of a telescope, its ability to make images bigger, is the least important of the three powers. Because the amount of detail that a telescope can discern is limited generally either by its resolving power or the seeing conditions, very high magnification does not necessarily show more detail. The magnifying power of a telescope equals the focal length of the primary mirror or lens divided by the focal length of the eyepiece. In other words, you can change the magnification of a telescope simply by changing the eyepiece. Radio telescopes of course don't have eyepieces, but they do have instruments that examine the radio waves focused by the telescope, and each such instrument would, in effect, have its own magnifying power.

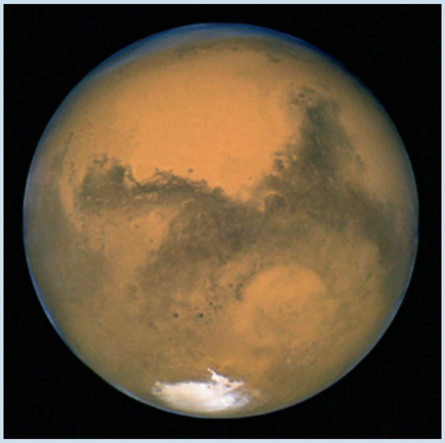
As was mentioned earlier, the two most important powers of the telescope, light-gathering power and resolving power, depend on the diameter of the telescope. This explains why astronomers describe telescopes by diameter and not by magnification. Astronomers will refer to a telescope as a 4-meter telescope or a 10-meter telescope, but they would never identify a research telescope as being, say, a 1000-power telescope. **Reasoning with Numbers 5-1** shows you how to calculate the powers of a telescope.

Resolution and Precision

What limits the detail you can see in an image? All images have limited resolution. You see this on your computer screen because images there are made up of picture elements, pixels. If your screen has large pixels, the resolution is low, and you can't see much detail. In an astronomical image, the size of a picture element is set by atmospheric seeing and by diffraction in the telescope. You can't see detail smaller than that resolution limit.

This limitation on detail in an image is related to the limited precision of any measurement. Imagine a zoologist trying to measure the length of a live snake by holding it along a meter stick. The wriggling snake is hard to hold, so it is hard to measure accurately. Also, meter sticks are usually not marked finer than millimeters. Both factors limit the precision of the measurement. If the zoologist said the snake was 43.28932 cm long, you might wonder if that is true. The resolution of this measurement technique does not justify the precision implied by all those digits.

Whenever you make a measurement you should ask yourself how precise that measurement can be. The precision of the measurement is limited by the resolution of the measurement technique, just as the amount of detail in a photograph is limited by its resolution. If you photograph a star, you would not be able to see details on its surface for a reason similar to why the zoologist can't measure the snake's length to high precision.



NASA, ESA, and The Hubble Heritage Team [STScI/AURA]

A high-resolution, visible-wavelength image of Mars made by the Hubble Space Telescope *reveals details such as mountains, craters, and the southern polar cap.*

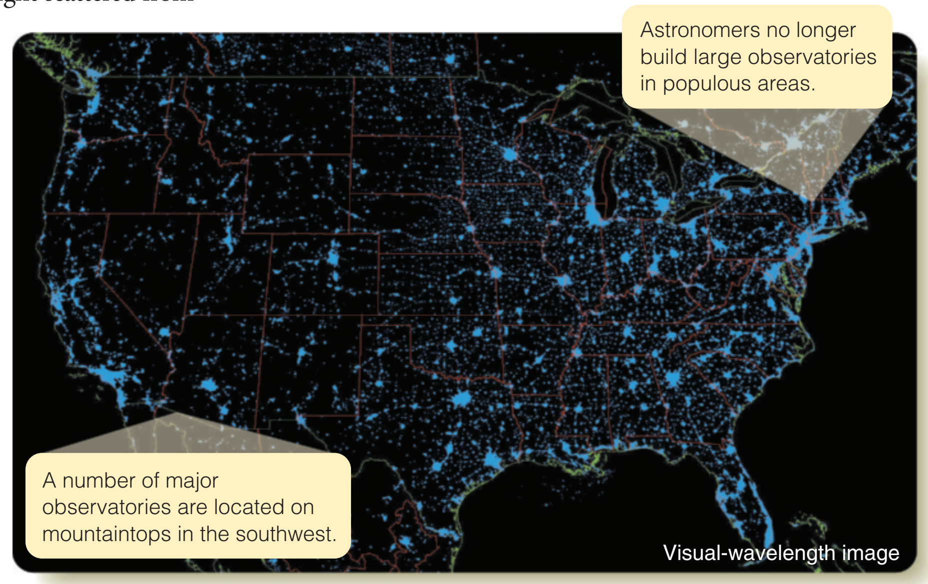
5-3 Observatories on Earth: Optical and Radio

The quest for light-gathering power and good resolution explains why nearly all the world's major observatories are located far from big cities and, especially in the case of optical telescopes, usually on top of mountains. Astronomers avoid cities because **light pollution**, the brightening of the night sky by light scattered from artificial outdoor lighting, can make it impossible to see faint objects (■ Figure 5-11). In fact, many residents of cities are unfamiliar with the beauty of the night sky because they can see only the brightest stars. Even far from cities, nature's own light pollution, the moon, is sometimes so bright it drowns out fainter objects, and astronomers are unable to perform certain types of observations during nights near full moon. On such nights, faint objects cannot be detected even with large telescopes at good locations.

Radio astronomers face a problem of radio interference comparable to visible light pollution. Weak radio waves from the cosmos are easily drowned out by human-made radio noise—everything from automobiles

■ **Figure 5-11**

This satellite view of the continental United States at night shows the light pollution, and energy waste, produced by outdoor lighting. Observatories are best located far from large cities.



© Cengage Learning 2014; image: NOAA

Reasoning with Numbers | 5-1

The Powers of a Telescope

The **light-gathering power** of a telescope is proportional to the area of the primary mirror or lens. A primary with a large area gathers a large amount of light. The area of a circular lens or mirror of diameter D is $\pi(D/2)^2$. To compare the relative light-gathering powers *(LGPs)* of two telescopes A and B, you can calculate the ratio of the areas of their primaries, which is the ratio of the diameters *(D)* squared:

$$\frac{LGP_A}{LGP_B} = \left(\frac{D_A}{D_B}\right)^2$$

Example A: Suppose you compare a 4-cm telescope with a 24-cm telescope. How much more light will the larger telescope gather? **Solution:**

$$\frac{LGP_{24}}{LGP_4} = \left(\frac{24}{4}\right)^2 = 6^2 = 36 \text{ times more light}$$

Example B: Your dark-adapted eye acts like a telescope with a diameter of about 0.8 cm, the maximum diameter of the pupil. How much more light can you gather if you use a 24-cm telescope, relative to your unaided eye? **Solution:**

$$\frac{LGP_{24}}{LGP_{eye}} = \left(\frac{24}{0.8}\right)^2 = 30^2 = 900 \text{ times more light}$$

The **resolving power** of a telescope is the angular distance between two stars that are just barely visible through the telescope as separate images. The resolving power α in arc seconds of a telescope with diameter D in meters that is collecting light of wavelength λ in meters equals:

$$\alpha(\text{arc seconds}) = 2.06 \times 10^5\left(\frac{\lambda}{D}\right)$$

The multiplication factor of 2.06×10^5 is the conversion between radians and arc seconds that you first learned in the small-angle formula (Chapter 3).

If the wavelength of light being studied is assumed to be 550 nm, in the middle of the visual band, then the above formula simplifies to:

$$\alpha(\text{arc seconds}) = \frac{0.113}{D}$$

Example C: What is the resolving power of a 10.0-cm (= 0.100 m, or about 4 in.) diameter telescope observing at visual wavelengths? **Solution:**

$$\alpha = 2.06 \times 10^5\left(\frac{550 \times 10^{-9}}{0.100}\right) = 1.13 \text{ arc seconds}$$

OR, equivalently,

$$\alpha = \frac{0.113}{0.100} = 1.13 \text{ arc seconds}$$

In other words, using a 10-cm (about 4-in.)-diameter telescope, you should be able to distinguish as separate points of light any pair of stars farther apart than about 1.1 arc seconds if the lenses are of good quality and if the seeing is good. Stars any closer together than that will be blurred together into a single image by the diffraction fringes.

Example D: The same resolving power formula can be applied to a radio telescope. What is the resolving power of a telescope with a dish (= primary mirror) diameter of 100 meters observing at a wavelength of 21.0 cm (= 0.210 m)? **Solution:**

$$\alpha = 2.06 \times 10^5\left(\frac{0.210}{100}\right) = 430. \text{ arc seconds}$$

Note how poor the resolution of even a large radio dish is, compared with optical telescopes. That resolution limit of 430 arc seconds corresponds to 1/4 the diameter of the full moon.

The **magnifying power** M of a telescope is the ratio of the focal length of the primary lens or mirror F_p divided by the focal length of the eyepiece F_e:

$$M = \left(\frac{F_p}{F_e}\right)$$

Example E: What is the magnification of a telescope with a primary mirror focal length of 80 cm if it is used with an eyepiece with a focal length of 0.50 cm? **Solution:** The magnification is 80 divided by 0.50, or 160 times.

Richard Wainscoat, Aina Kai LLC

Figure 5-12

Optical, infrared, and microwave radio telescopes at 4,200 meters (nearly 14,000 feet) above sea level on Mauna Kea in Hawai'i. The high altitude, low atmospheric moisture, lack of nearby large cities, and location near the equator make this mountain one of the best places on Earth to build an observatory.

with faulty spark plugs to poorly designed communication systems. A few narrow radio bands are reserved for astronomy research, but even those are often contaminated by stray signals. To avoid that noise and have the radio equivalent of a dark sky, astronomers locate radio telescopes as far from civilization as possible. Hidden in mountain valleys or in remote deserts, they are able to study the universe protected from humanity's radio output.

As you have already learned, astronomers prefer to put optical telescopes on high mountains for several reasons. For the best seeing, astronomers carefully select mountains where the air flow is measured to be smooth and not turbulent. Also, the air at high altitude is thin, dry, and more transparent, which is important not only for optical telescopes but also for some types of radio telescopes. Building an observatory on top of a remote, high mountain is difficult and expensive, as you can imagine from Figure 5-12, but the dark sky, good seeing, and transparent atmosphere make it worth the effort.

Modern Optical Telescopes

For most of the 20th century, astronomers faced a serious limitation on the size of astronomical telescopes. Telescope mirrors were made thick to avoid bending that would distort the reflecting surface, but those thick mirrors were heavy. The 5-m (200-in.) mirror on Mount Palomar weighs 14.5 tons. Those old-fashioned telescopes were big, heavy, and expensive.

Today's astronomers have solved these problems in a number of ways. Read **Modern Optical Telescopes** on pages 88–89 and notice three important points about telescope design and ten new terms that describe optical telescopes and their operation:

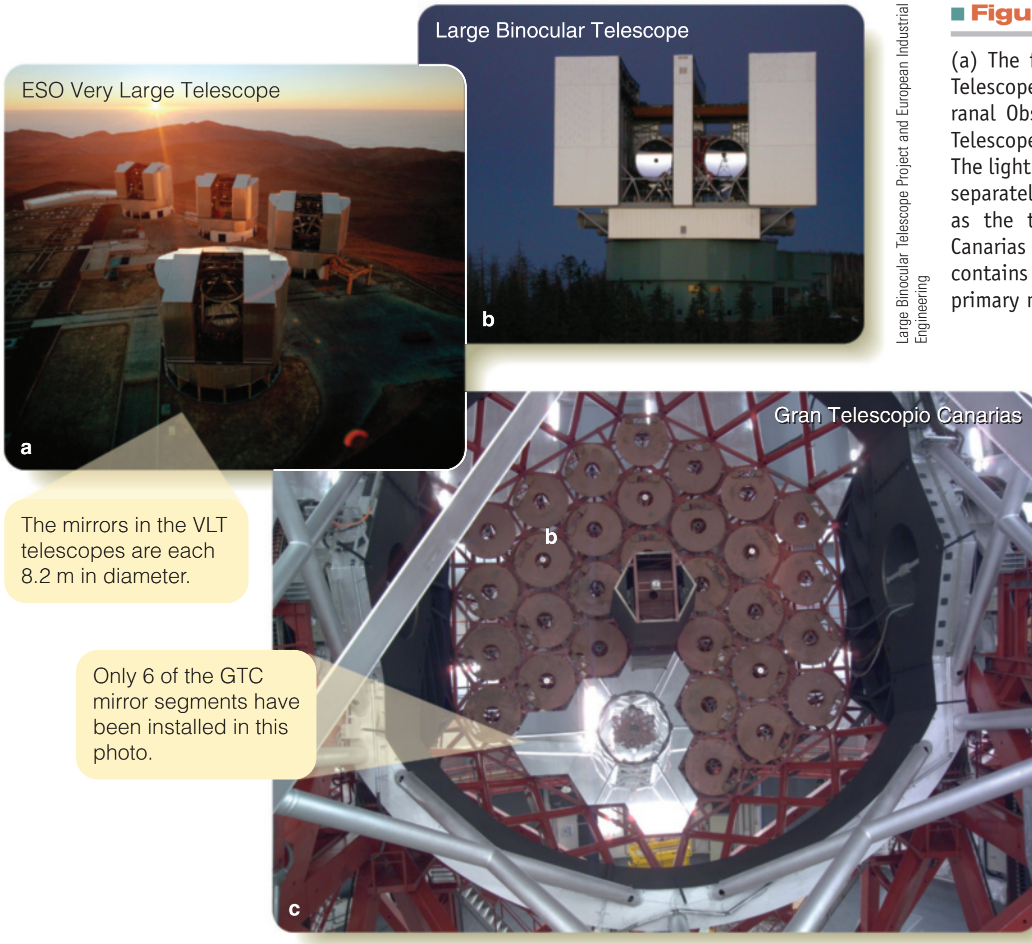
1. Conventional-design reflecting telescopes use large, solid, heavy mirrors to focus starlight to a *prime focus,* or, by using a *secondary mirror,* to a *Cassegrain focus.* Other telescopes have a *Newtonian focus* or a *Schmidt-Cassegrain focus.*

2. Telescopes must have a *sidereal drive* to follow the stars ("sidereal," pronounced *sih-DARE-ee-al,* refers to stars). An *equatorial mount* with motion around a *polar axis* is the conventional way to provide that motion. Today, astronomers can build simpler, lighter-weight telescopes on *alt-azimuth mounts* that depend on computers to move the telescope so that it follows the apparent motion of stars as Earth rotates without having an equatorial mount and polar axis.

3. *Active optics,* computer control of the shape of a telescope's main mirrors, allows the use of thin, lightweight mirrors—either "floppy" mirrors or segmented mirrors. Reducing the weight of the mirror reduces the weight of the rest of the telescope, making it stronger and less expensive. Also, thin mirrors cool and reach a stable shape faster at nightfall, producing better images during most of the night.

Modern engineering techniques and high-speed computers have allowed astronomers to build and use new, giant telescopes with unique designs. A few are shown in Figure 5-13. The European Southern Observatory has built the Very Large Telescope (VLT) in the foothills of the Andes Mountains in northern Chile. The VLT actually consists of four telescopes, each with a computer-controlled mirror 8.2 m in diameter and only 17.5 cm (6.9 in.) thick. Italian and American astronomers have built the Large Binocular Telescope (LBT) on Mount Graham in Arizona. The LBT carries a pair of 8.4-m mirrors on a single mount. The twin Keck telescopes on Mauna Kea in Hawai'i have primary mirrors 10 meters (400 in., more than 33 ft) in diameter that are each made of 36 individually controlled hexagonal segments. The Gran Telescopio Canarias (GTC), located atop a volcanic peak in the Canary Islands, carries a segmented mirror 10.4 meters in diameter and is, at the time of this writing, the largest single telescope in the world.

Other giant telescopes are being planned with segmented mirrors or with multiple mirrors (Figure 5-14). The Giant Magellan Telescope (GMT) will carry seven thin mirrors, each 8.4 meters in diameter, on a single mounting. It will be located in Chile and will have the light-gathering power of a single 24.5-m telescope. The Thirty Meter Telescope (TMT), now under development by American astronomers, is planned to have a mirror up to 30 meters (100 ft) in diameter comprised of 492 hexagonal segments and will be located on Mauna Kea in Hawai'i. An international team is designing the European Extremely Large Telescope (E-ELT) to carry 906 segments, making up a mirror 42 meters in diameter. The E-ELT will be built



ESO

Large Binocular Telescope Project and European Industrial Engineering

Gara Mora, IAC

© Cengage Learning 2014

■ Figure 5-13

(a) The four telescopes of the European Very Large Telescope (VLT) are housed in separate domes at Paranal Observatory in Chile. (b) The Large Binocular Telescope (LBT) in Arizona carries two 8.4-m mirrors. The light gathered by the two mirrors can be analyzed separately or combined. The entire building rotates as the telescope moves. (c) The Gran Telescopio Canarias (GTC) on La Palma in the Canary Islands contains 36 hexagonal mirror segments in its 10.4-m primary mirror.

on Cerro Armazones, a mountain in the Atacama desert of Chile. Other very large telescopes have been proposed with estimated completion dates of 2017 or later.

Ground-based telescopes usually are used by astronomers and technicians working in a control room in the same building or nearby, but some telescopes are now operated by astronomers thousands of miles from the observatory. Other telescopes are fully automated and operate without direct human supervision. That has made possible huge surveys of the sky in which millions of objects are observed. The Sloan Digital Sky Survey (SDSS), for example, mapped the Northern Hemisphere sky, measuring the position and brightness of 100 million stars and galaxies at five ultraviolet, optical, and infrared wavelengths. The Two-Micron All Sky Survey (2MASS) mapped the entire sky at three near-infrared wavelengths. More surveys are being made or planned at other wavelengths. For example, the Large Synoptic Survey Telescope (LSST) is envisioned as an 8.4-m telescope with a 3.2-billion-pixel CCD camera that will be able to record the brightness at near-UV, visual, and near-IR wavelengths of every object brighter than magnitude 24.5 in one hemisphere of the sky every three nights. Astronomers will study those databases for decades to come.

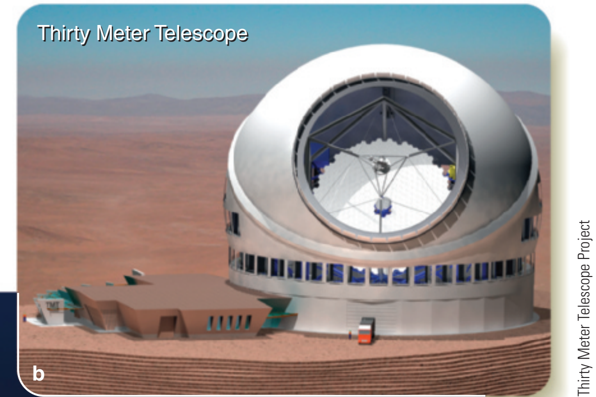
Modern Radio Telescopes

The dish reflector of a radio telescope, like the mirror of a reflecting telescope, collects and focuses radiation. Although a radio telescope's dish may be tens or hundreds of meters in diameter, the receiver antenna may be as small as your hand. Like the antenna on a TV set, its only function is to absorb the radio energy collected by the dish. Because radio wavelengths are in the range of a few millimeters to a few tens of meters, the dish only needs to be shaped to that level of accuracy, much less smooth than a good optical mirror. In fact, wire mesh works well as a mirror for all but the shortest-wavelength radio waves.

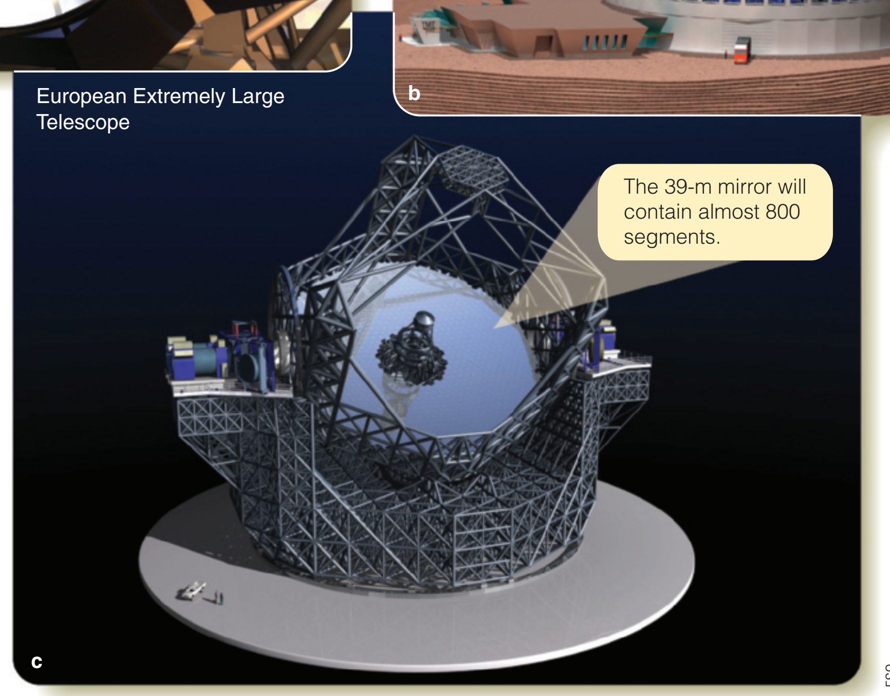
The largest radio dish in the world is 305 m (1000 ft) in diameter. Such a large dish can't be supported easily, so it is built into a mountain valley in Arecibo, Puerto Rico. The reflecting dish is a thin metallic surface supported above the valley floor by cables attached near the rim, and the antenna platform hangs above the dish on cables from towers built on three mountain peaks that surround the valley (■ Figure 5-15). By moving the antenna above the dish, radio astronomers can point the telescope at any object that passes within 20 degrees of the zenith as Earth rotates. Since completion in 1963, the Arecibo telescope has been an international center of radio astronomy research. China is beginning construction of a 500-m-diameter radio telescope named FAST, located in a mountain-ringed valley like Arecibo's.

A radio astronomer works under two disadvantages relative to optical astronomers: poor resolution and low signal intensity. Recall that the resolving power of a telescope depends on the diameter of the primary lens or mirror and also on the wavelength of the radiation. At very long wavelengths like those

Carnegie Observatories



Thirty Meter Telescope Project



ESO

Figure 5-14

Three proposed new telescopes. (a) The Giant Magellan Telescope (GMT) will have the resolving power of a telescope 24.5 m in diameter when it is finished in about 2020. (b) The Thirty Meter Telescope (TMT) is planned to occupy a dome specially designed to be as small as possible given the telescope size. (c) Like nearly all of the newest large telescopes, the European Extremely Large Telescope (E-ELT) will have an alt-azimuth mount. Note the car and people at lower left for scale.

© Cengage Learning 2014

of radio waves, the diffraction fringes are relatively large. This means that images or maps from individual radio telescopes generally don't show such fine details as are seen in optical images.

The second handicap radio astronomers face is the low intensity of the radio signals. You learned earlier that the energy of a photon depends on its wavelength. Photons of radio energy have such long wavelengths that their individual energies are quite low. The cosmic radio signals arriving on Earth are astonishingly weak—a million to a billion times weaker than the signal from a commercial radio station. To get detectable signals focused on the antenna, radio astronomers must build large collecting areas either as single large dishes or by combining

David Parker

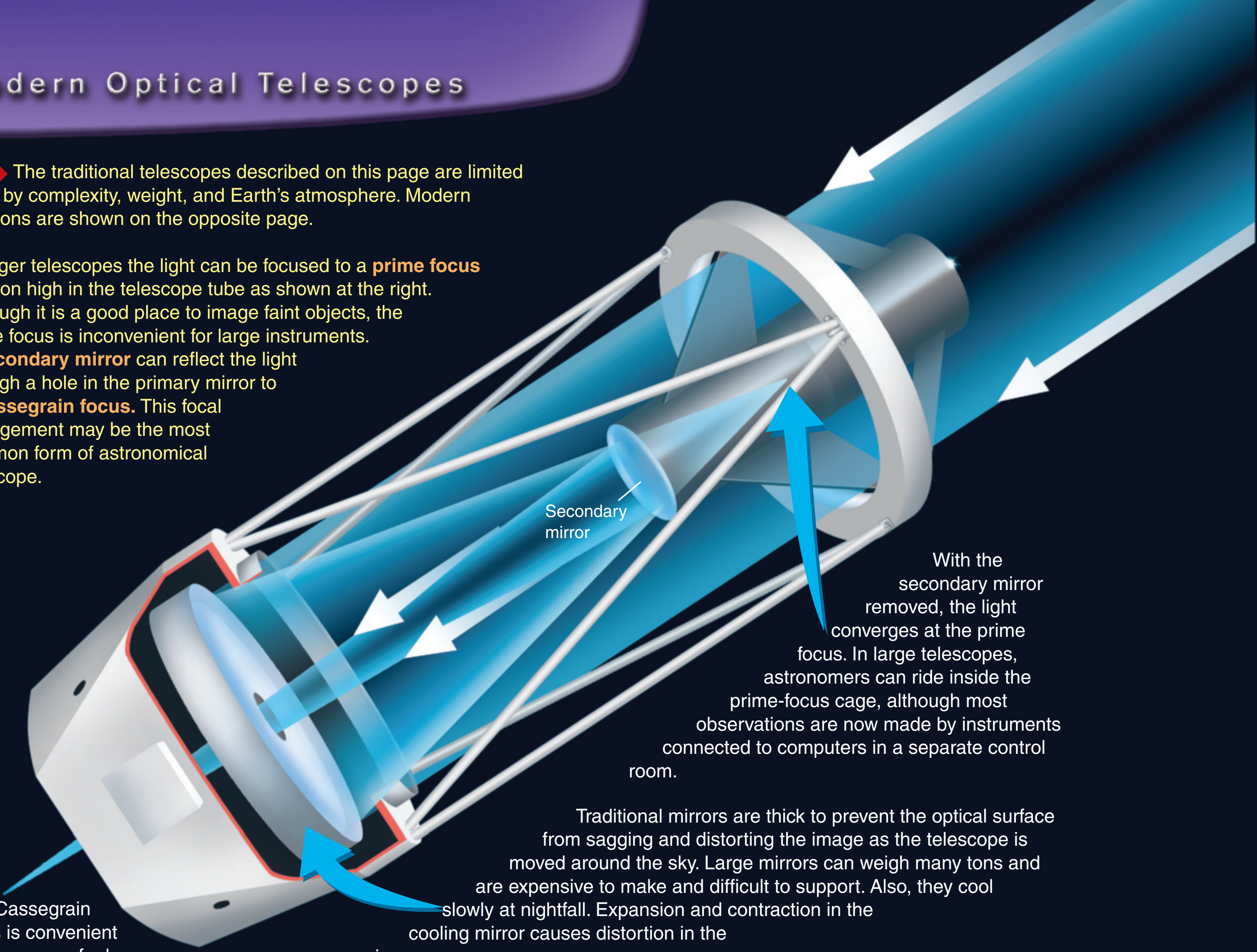
Figure 5-15

The 305-m (1000-ft) radio telescope in Arecibo, Puerto Rico, is nestled in a naturally bowl-shaped valley. The receiver platform is suspended over the dish. A consortium led by SRI International and Universities Space Research Association manages Arecibo Observatory for the National Science Foundation.

Modern Optical Telescopes

1 The traditional telescopes described on this page are limited by complexity, weight, and Earth's atmosphere. Modern solutions are shown on the opposite page.

In larger telescopes the light can be focused to a **prime focus** position high in the telescope tube as shown at the right. Although it is a good place to image faint objects, the prime focus is inconvenient for large instruments. A **secondary mirror** can reflect the light through a hole in the primary mirror to a **Cassegrain focus.** This focal arrangement may be the most common form of astronomical telescope.

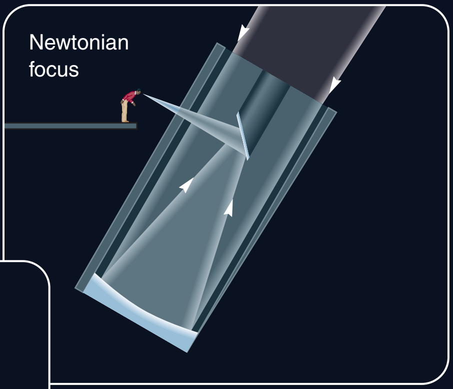


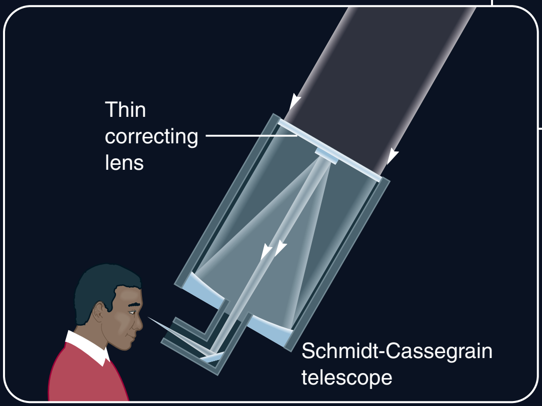
With the secondary mirror removed, the light converges at the prime focus. In large telescopes, astronomers can ride inside the prime-focus cage, although most observations are now made by instruments connected to computers in a separate control room.

Traditional mirrors are thick to prevent the optical surface from sagging and distorting the image as the telescope is moved around the sky. Large mirrors can weigh many tons and are expensive to make and difficult to support. Also, they cool slowly at nightfall. Expansion and contraction in the cooling mirror causes distortion in the images.

The Cassegrain focus is convenient and has room for large instruments.

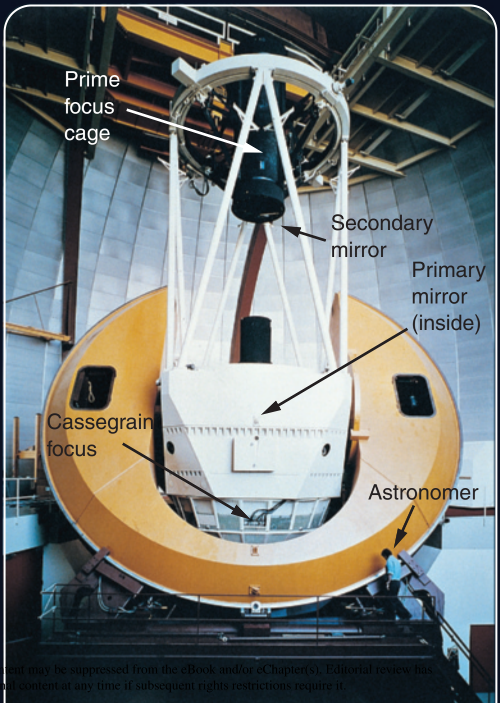
1a Smaller telescopes are often found with a **Newtonian focus,** the arrangement that Isaac Newton used in his first reflecting telescope. The Newtonian focus is inconvenient for large telescopes as shown at right.





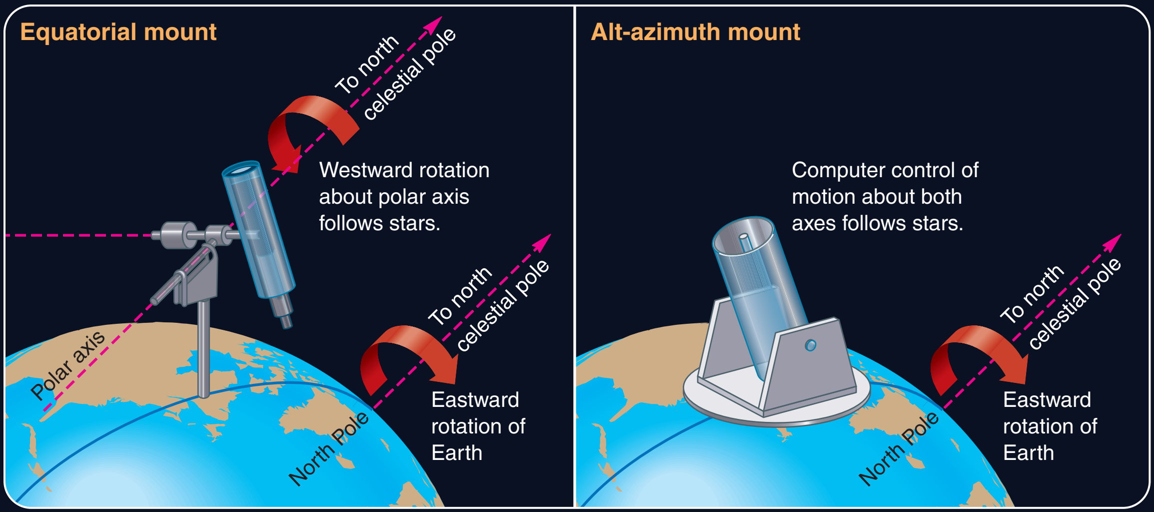
1b Many small telescopes such as the one on your left use a **Schmidt-Cassegrain focus.** A thin correcting plate improves the image but is too slightly curved to introduce serious chromatic aberration.

1c Shown below, the 4-meter Mayall Telescope at Kitt Peak National Observatory in Arizona can be used at either the prime focus or the Cassegrain focus. Note the human figure at lower right.



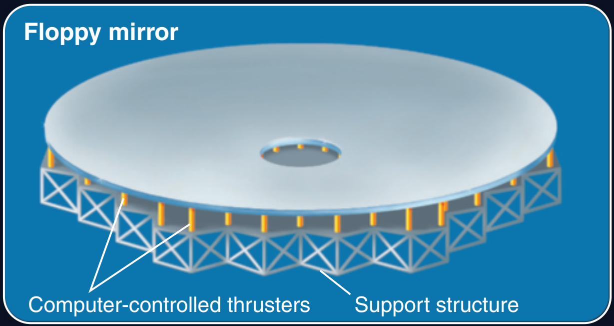
AURA/NOAO/NSF

Line art on this page © Cengage Learning 2014



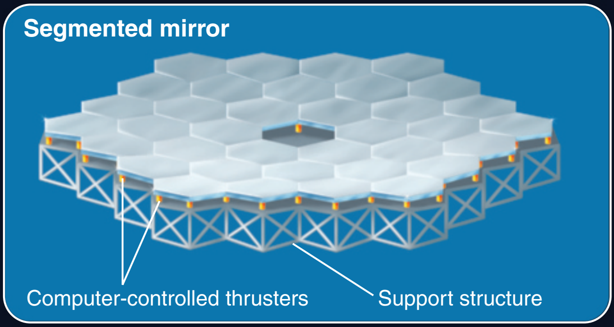
2 Telescope mountings must contain a **sidereal drive** to move smoothly westward and counter the eastward rotation of Earth. The traditional **equatorial mount** (far left) has a **polar axis** parallel to Earth's axis, but the modern **alt-azimuth mount** (near left) moves like a cannon — up and down and left to right. Such mountings are simpler to build but need computer control to follow the stars.

3 Unlike traditional thick mirrors, thin mirrors, sometimes called floppy mirrors as shown at right, weigh less and require less massive support structures. Also, they cool rapidly at nightfall and there is less distortion from uneven expansion and contraction.

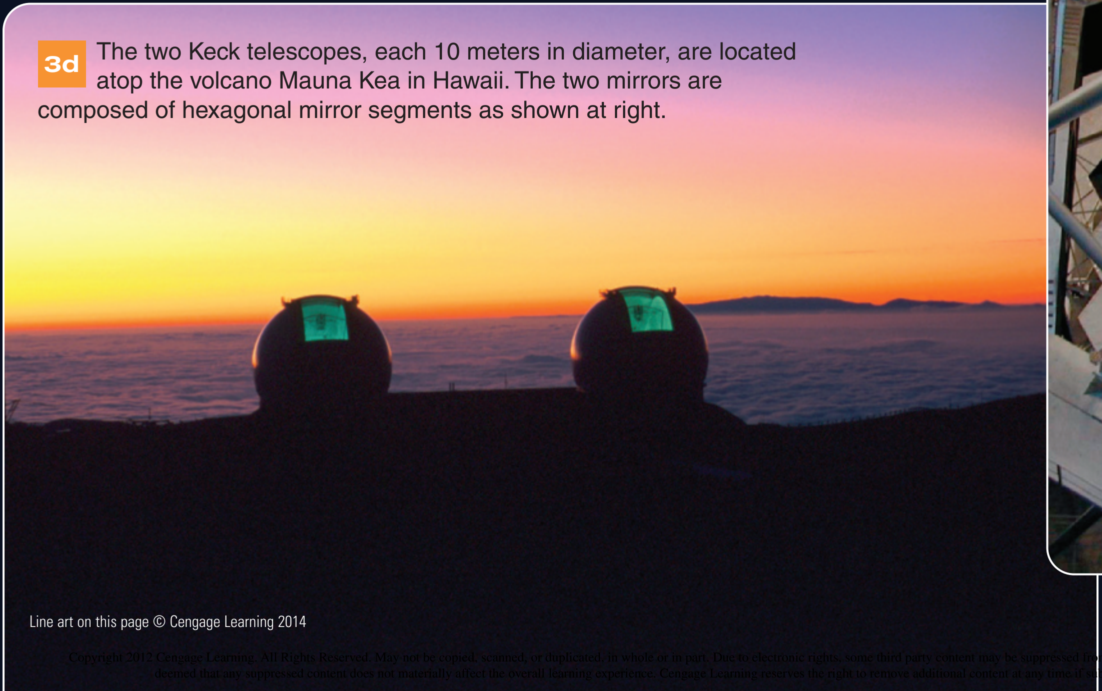


3a Grinding a large mirror may remove tons of glass and take months, but new techniques speed the process. Some large mirrors are cast in a rotating oven that causes the molten glass to flow to form a concave upper surface. Grinding and polishing such a preformed mirror is much less time consuming.

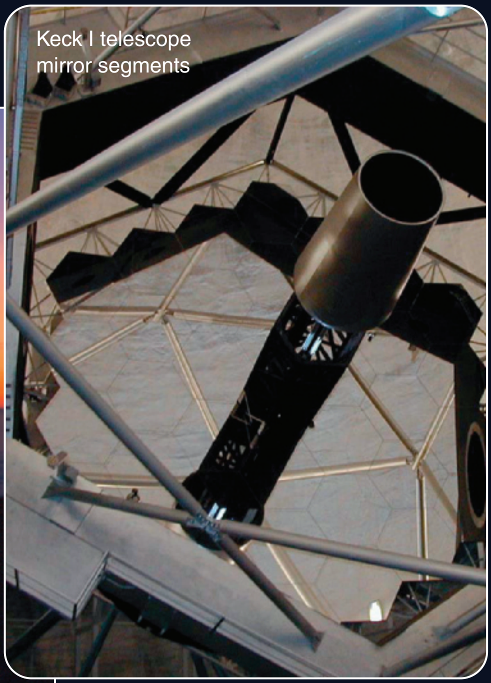
3b Mirrors made of segments are economical because the segments can be made separately. The resulting mirror weighs less and cools rapidly. See image at right.



3c Both floppy mirrors and segmented mirrors sag under their own weight. Their optical shape must be controlled by computer-driven thrusters under the mirror in what is called **active optics.**



3d The two Keck telescopes, each 10 meters in diameter, are located atop the volcano Mauna Kea in Hawaii. The two mirrors are composed of hexagonal mirror segments as shown at right.



W.M. Keck Observatory

arrays of smaller dishes. Even then, because the radio energy from celestial objects is so weak, it must be strongly amplified before it can be measured and recorded.

SCIENTIFIC ARGUMENT

Why do astronomers build optical observatories at the tops of mountains? Measurement accuracy is so important that scientists often take extreme steps to gather accurate information. It certainly isn't easy to build a large, complicated, and delicate optical telescope at the top of a high mountain, but it is worth the effort. A telescope on top of a high mountain is above the densest part of Earth's atmosphere. There is less air to dim the incoming light. Even more important, turbulence of thin air on a mountaintop is less able to disturb light waves than that of thick air, so the seeing is better. The resolving power of a large optical telescope on Earth's surface is set by atmospheric seeing rather than by the telescope's diffraction. It really is worth the trouble to build telescopes on the peaks of high mountains.

Astronomers also put radio observatories in special locations. Extend your argument to that type of astronomical observations. **What considerations do astronomers make in choosing the location of a new radio telescope?**

5-4 Airborne and Space Observatories

Ground-based telescope performance is limited by Earth's atmospheric turbulence and transparency. As you will learn in the next section, there are sophisticated techniques that partly compensate for atmospheric seeing, but a telescope in space has no such problem, and its resolution is defined only by diffraction.

Also, a telescope on the ground must look through one of the two open atmospheric "windows" in the visible or radio parts of the electromagnetic spectrum described earlier in this chapter. Most types of electromagnetic radiation arriving here from the universe—infrared, ultraviolet, X-rays, and gamma rays—do not reach Earth's surface because they are partly or completely absorbed by Earth's atmosphere. To operate at those blocked wavelengths, telescopes must go to high altitudes or into space. As you will learn in the next few chapters, objects that are cooler than stars, such as stars that are forming, emit mostly infrared and radio radiation and relatively little visible light. In contrast, cosmic catastrophes such as exploding stars make mostly gamma rays and X-rays. Combining information from as wide a variety of wavelengths as possible allows astronomers to gain a more complete understanding of the universe.

The *Hubble Space Telescope*

Named after Edwin Hubble, the astronomer who discovered the expansion of the universe, the *Hubble Space Telescope* is the most successful observatory in history (■ Figure 5-16a). It was launched in 1990 and contains a 2.4-m (95-in.) mirror plus instruments with which it can observe visible, near-ultraviolet, and near-infrared light. It is controlled from a research center on Earth and observes almost continuously. Nevertheless, the telescope has time to complete only a fraction of the projects proposed by astronomers from around the world.

Most of the observations *Hubble* makes are at visual wavelengths, so its greatest advantage in being located above Earth's atmosphere is the lack of seeing distortion. It therefore can detect fine detail, and, because it concentrates light into sharp images, it can detect extremely faint objects.

The *Hubble* telescope is as big as a city bus. It has been visited a number of times by the space shuttle so that astronauts could service its components and install new cameras and other instruments. Thanks to the work of the space shuttle astronaut crew that visited in 2009 and refurbished the telescope's instruments, batteries, and gyroscopes, *Hubble* will almost certainly last until it is replaced by the *James Webb Space Telescope (JWST),* which is expected to be launched into solar orbit in about the year 2018. The *James Webb* telescope will carry a cluster of beryllium segments that will open in space to form a 6.5-m (256-in.) mirror (Figure 5-16b).

Infrared Astronomy from Aircraft and Spacecraft

In addition to the atmospheric windows at visual and radio wavelengths you have already learned about, there are also a few narrow windows at short infrared wavelengths accessible from the ground, especially from high mountains such as Mauna Kea (Figure 5-12). However, most infrared wavelengths are blocked, especially by water vapor absorption. Also, Earth's atmosphere itself glows in the far-infrared. Observations at very long infrared wavelengths can be made only from telescopes carried to high altitudes by aircraft or balloons, or launched entirely out of the atmosphere onboard spacecraft. (Notice that the reasons to get an infrared telescope above the atmosphere are somewhat different from the reasons to send an optical telescope into space.)

In the 1960s through the 1990s NASA developed a series of infrared observatories with telescopes carried above Earth's atmospheric water vapor by jet aircraft. The modern successor to those airborne facilities, the *Stratospheric Observatory for Infrared Astronomy (SOFIA),* is a 2.5-meter (100-in.) telescope looking out a roll-back opening in the left side of the fuselage of an extensively modified Boeing 747SP aircraft (■ Figure 5-17a). Astronomers used SOFIA for more than 200 hours of scientific observations during 2011, its first full year of operations.

Telescopes with the most sensitive long-wavelength infrared detectors need to get completely above Earth's absorbing atmosphere and also must be protected from heat. They have limited

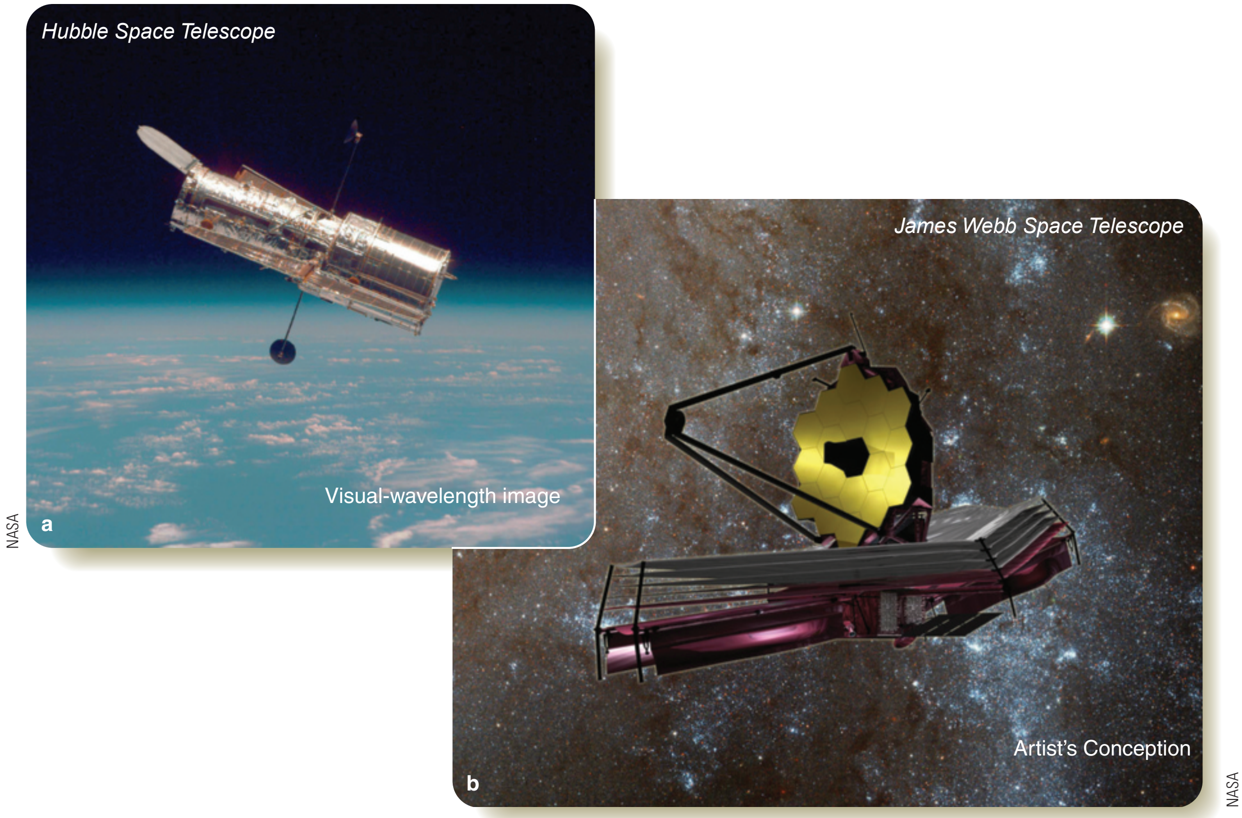


Figure 5-16

(a) The *Hubble Space Telescope* orbits Earth at an average of 570 km (355 mi) above the surface. In this image the telescope is viewing toward the upper left. (b) An artist's conception of *Hubble*'s eventual successor, the *James Webb Space Telescope (JWST)*. *JWST* will be located in solar orbit 1 million miles from Earth and will not have an enclosing tube, thus resembling a radio dish more than a conventional optical telescope. It will observe the universe from behind a multilayered sun screen larger than a tennis court.

lifetimes because they carry coolant to chill their optics to near absolute zero temperature (−273°C, or −460°F), and the coolant eventually runs out. The European Space Agency's *Herschel* 3-m infrared space telescope, pictured in Figure 5-17c, was launched (together with the smaller *Planck* millimeter-wavelength space telescope) into solar orbit in 2009. During its three-year mission, *Herschel* made important discoveries concerning star formation, planet formation, and chemistry in interstellar clouds.

High-Energy Astronomy

Like infrared-emitting objects, gamma-ray, X-ray and ultraviolet sources in the universe are difficult to observe because the telescopes must be located high in Earth's atmosphere or in space. Some high-energy astronomy satellites have been general-purpose telescopes that can observe many different kinds of objects, for example, *ROSAT,* an X-ray observatory developed by a multinational consortium of European astronomers. In contrast, some space telescopes are designed to study a single problem or a single object. For example, the Japanese satellite *Hinode* (pronounced, *hee-no-day*) studies the sun continuously at visual, ultraviolet, and X-ray wavelengths.

The largest X-ray telescope to date is the *Chandra X-ray Observatory (CXO). Chandra* operates in an orbit that extends a third of the way to the moon. Focusing X-rays is difficult because they penetrate into most mirrors, so astronomers devised cylindrical mirrors in which the X-rays reflect at shallow angles from the polished inside of the cylinders to form images on X-ray detectors, as shown in Figure 5-18. The *Chandra* observatory has made important discoveries about everything from star formation to monster black holes in distant galaxies that will be described in later chapters.

One of the first gamma-ray observatories was the *Compton Gamma Ray Observatory,* launched in 1991. It mapped the entire sky at gamma-ray wavelengths. The European *INTEGRAL (International Gamma-Ray Astrophysics Laboratory)* satellite was launched in 2002 and has been very productive in the study of violent eruptions of stars and black holes. The *GLAST (Gamma-Ray Large Area Space Telescope),* launched in 2008, is capable of mapping large areas of the sky to high sensitivity.

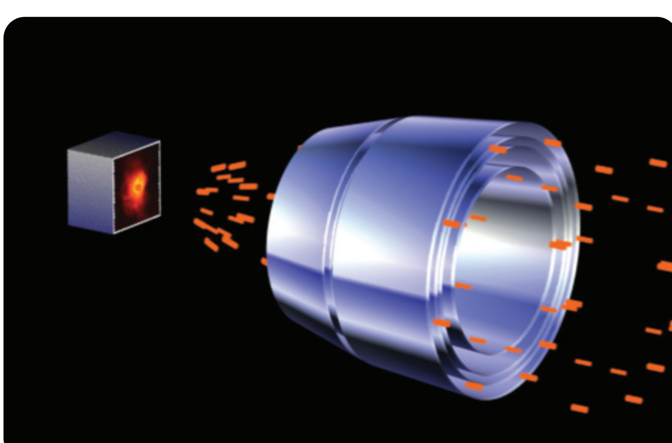
NASA

Anthony Wesley; NASA/DLR/USRA/DSI/FORCAST Team

ESA/D. Ducros

Figure 5-17

(a) *SOFIA, the Stratospheric Observatory for Infrared Astronomy,* a joint project of NASA and the German Aerospace Center (DLR), will fly at altitudes up to 14 km (45,000 ft), where it will be able to collect infrared radiation at far infrared wavelengths that are unobservable even from high mountaintops. (b) 1Infrared image of the planet Jupiter at wavelengths of 5.4, 24, and 37 microns (right) made during *SOFIA's* "First Light" flight in May 2010 compared with a visual-wavelength image (left). The white stripe in the infrared image is a region of relatively transparent clouds through which the warm interior of the planet can be seen. (c) Artist's conception of the *Herschel* infrared space telescope, carrying a 3-m mirror and instruments cooled almost to about 70 K.



NASA/SAO/CXC

Figure 5-18

X-rays that hit a mirror at grazing angles are reflected like a pebble skipping across a pond. Thus, X-ray telescope mirrors like the ones in *Chandra* are shaped like barrels rather than dishes.

Modern astronomy has come to depend on observations that cover the entire electromagnetic spectrum. More orbiting space telescopes are planned that will be even more versatile and sensitive than the ones operating now.

5-5 Astronomical Instruments and Techniques

Just looking through a telescope doesn't tell you much. A star looks like a point of light. A planet looks like a little disk. A galaxy looks like a hazy patch. To use a research telescope to learn about the universe, you need to carefully analyze the light the telescope gathers. Special instruments attached to the telescope make that possible.

Imaging Systems and Photometers

The **photographic plate** was the first device used by astronomers to record images of celestial objects. Photographic plates can detect faint objects in long time exposures and can be stored for later analysis. Brightness of objects imaged on a photographic plate can be measured, but only with a lot of hard work that yields just moderate precision. Astronomers also build **photometers**, sensitive light meters used to measure the brightness of individual objects very precisely.

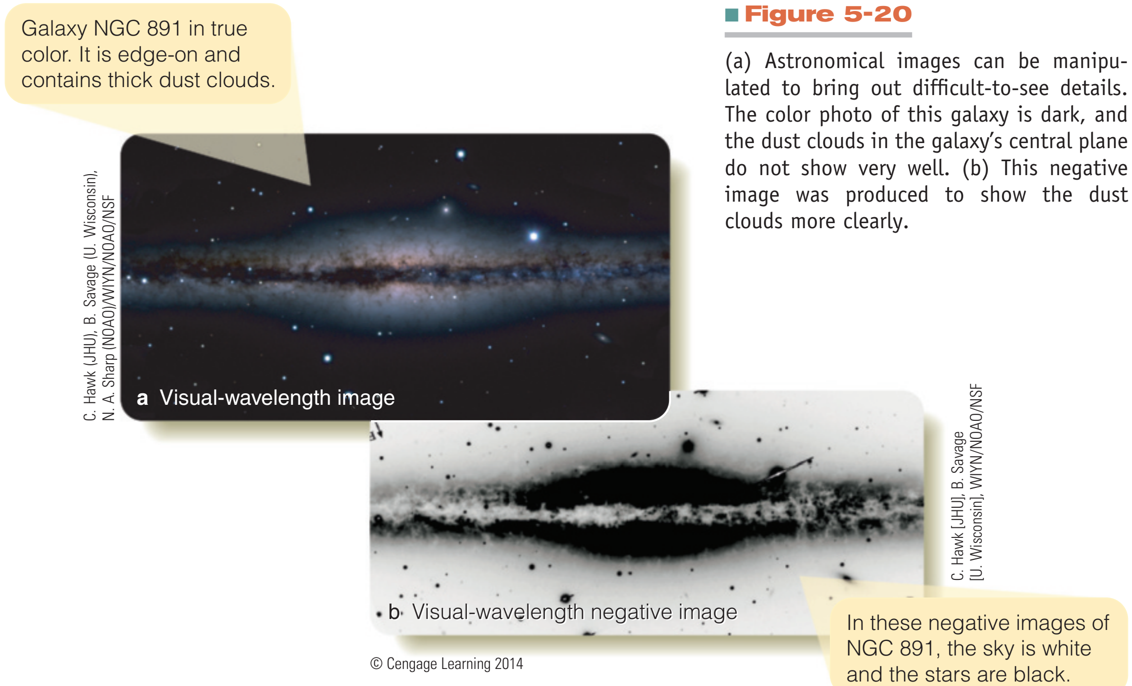
Present-day astronomers use **charge-coupled devices (CCDs)** as both image-recording devices and photometers. A CCD is a specialized computer chip containing millions of microscopic light detectors arranged in an array about the size of a postage stamp. CCD chips have replaced photographic plates because they have some dramatic advantages. CCDs are much more sensitive than photographic plates and can detect both bright and faint objects in a single exposure. Also, CCD images are **digitized**, meaning converted to numerical data, and thus can be stored in a computer's memory for later analysis. Although astronomy research CCDs are extremely sensitive and therefore expensive, less sophisticated CCDs are now part of everyday life. They are used in digital cameras (both still and video) as well as in cell phone cameras. Infrared astronomers use **array detectors** that are similar in operation to optical CCDs. At other wavelengths, photometers are still used for measuring brightness of celestial objects. Array detectors and photometers generally must be cooled to operate properly (■ Figure 5-19).

© Cengage Learning 2014; image: Kris Koenig/Coast Learning Systems

■ **Figure 5-19**

Astronomical cameras with array detectors must be cooled to low temperatures to operate properly, and that is especially true for infrared cameras.

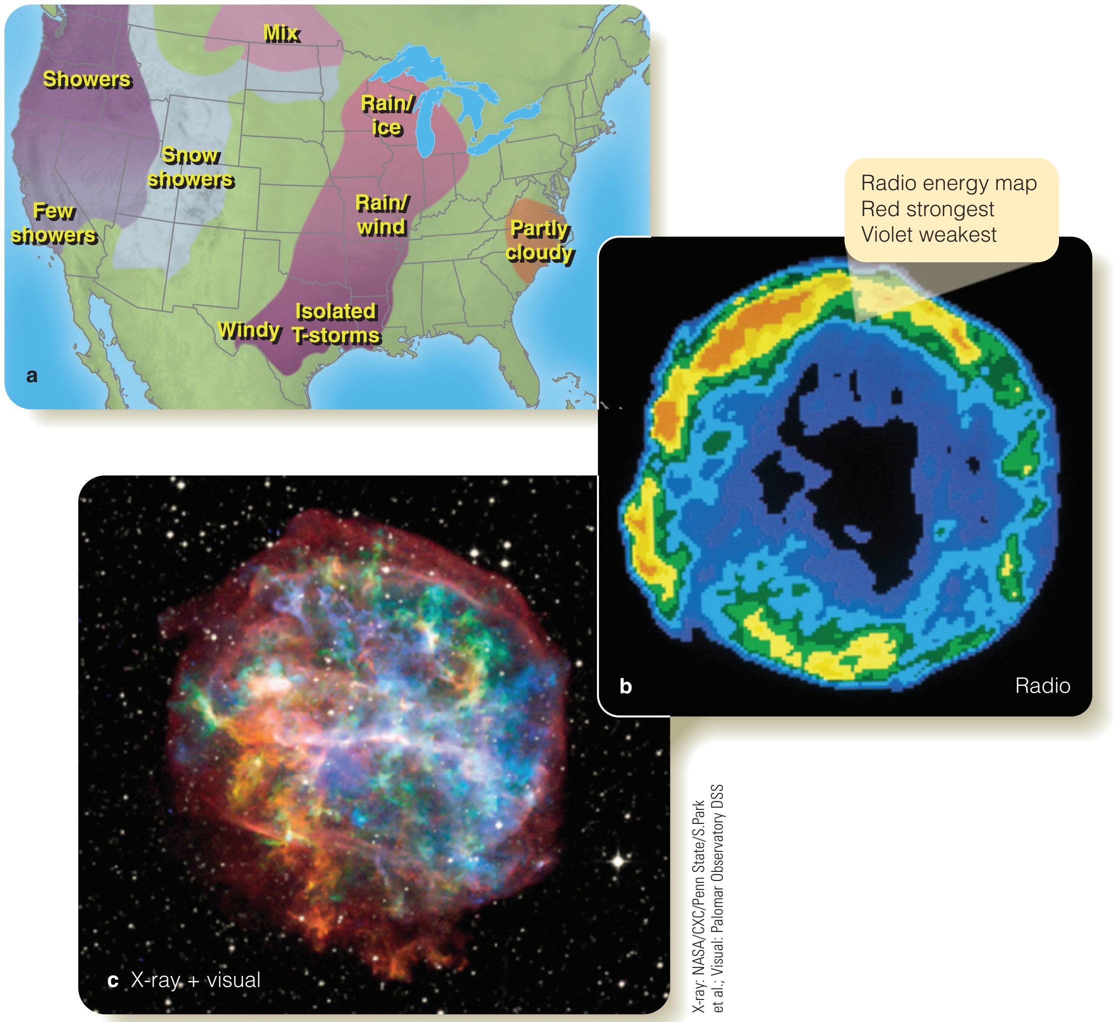
The digital data representing an image from a CCD or other array detector are easy to manipulate to bring out details that would not otherwise be visible. For example, astronomical images are often reproduced as negatives, with the sky white and the stars dark. That makes the faint parts of the image easier to see (■ Figure 5-20). Astronomers also manipulate images to produce **representational-color** images (also known as **false-color** images) in which the colors represent different aspects of the object such as intensity, rather than visual color. For example, because humans can't see radio waves, astronomers must convert radio data into something perceptible. One way is to measure the strength of the radio signal at various places in the sky and produce a representational-color map in which each color marks areas of similar radio intensity. Compare such a map to a weather map in which the different colors mark areas forecast to have different types and amounts of precipitation (■ Figure 5-21a). Representational-color images and maps are very commonly used in nonoptical astronomy (Figures 5-21b and 5-21c).



© Cengage Learning 2014

■ **Figure 5-20**

(a) Astronomical images can be manipulated to bring out difficult-to-see details. The color photo of this galaxy is dark, and the dust clouds in the galaxy's central plane do not show very well. (b) This negative image was produced to show the dust clouds more clearly.



© Cengage Learning 2014

Figure 5-21

(a) A typical weather map uses contours with added color to show which areas are likely to receive precipitation, and what type. (b) A radio image of Tycho's supernova remnant, the expanding shell of gas produced by the explosion of a star visible on Earth in the year 1572. This image's representational-color code shows intensity of radio radiation at one wavelength. (c) Composite X-ray and visual-wavelength image of supernova remnant G292.0+1.8. Colors are used to represent a range of X-ray wavelengths and white signifies emission at visual wavelengths. The shape and complex structure of the nebula hint at the history of the star before it exploded.

Spectrographs

To analyze light in detail, astronomers spread out the light according to wavelength (color), a function performed by a **spectrograph**. You can understand how this instrument works if you imagine repeating an experiment performed by Isaac Newton in 1666. Newton bored a small hole in the window shutter of his room to admit a thin beam of sunlight. When he placed a prism in the beam, it spread the light into a beautiful spectrum that splashed across his wall. From that and related experiments, Newton concluded that white light is made of a mixture of all the colors.

As you learned earlier in the context of refracting telescopes' chromatic aberration, light passing from one medium such as air into another medium such as glass has its path bent at an angle that depends on its wavelength. Thus, the white light entering the prism is spread into a spectrum leaving the prism (■ Figure 5-22). You can build a simple spectrograph by using a narrow opening to define the incoming light beam, a prism to spread the light into its component colors, and a lens to guide the light into a camera.

Almost all modern spectrographs use a **grating** rather than a prism. A grating is a piece of glass or metal with thousands of microscopic parallel grooves scribed onto its surface. Different wavelengths of light reflect from the grating at slightly different angles, so white light is spread into a spectrum. You have probably noticed this effect when you look at the closely spaced lines etched onto a music CD or video DVD; as you tip the disk, different colors flash across its surface. A modern spectrograph can be built using a high-quality grating to separate light by wavelength, plus a CCD detector to record the resulting spectrum.

The spectrum of an astronomical object can contain hundreds of **spectral lines**—dark or bright lines that cross the spectrum at specific wavelengths. You will learn in Chapters 6, 7, and 13 that spectra of the sun and stars contain hundreds of dark spectral lines produced by the atoms and molecules in the atmospheric gases of those objects. Because scientists understand the details of how light interacts with matter, a spectrum carries a tremendous amount of information. That makes a spectrograph the astronomer's most powerful instrument. In the next chapter, you will learn more about the information astronomers can extract from a spectrum. Some astronomers say, "We don't know anything about an object until we get a spectrum," and that is only a slight exaggeration.

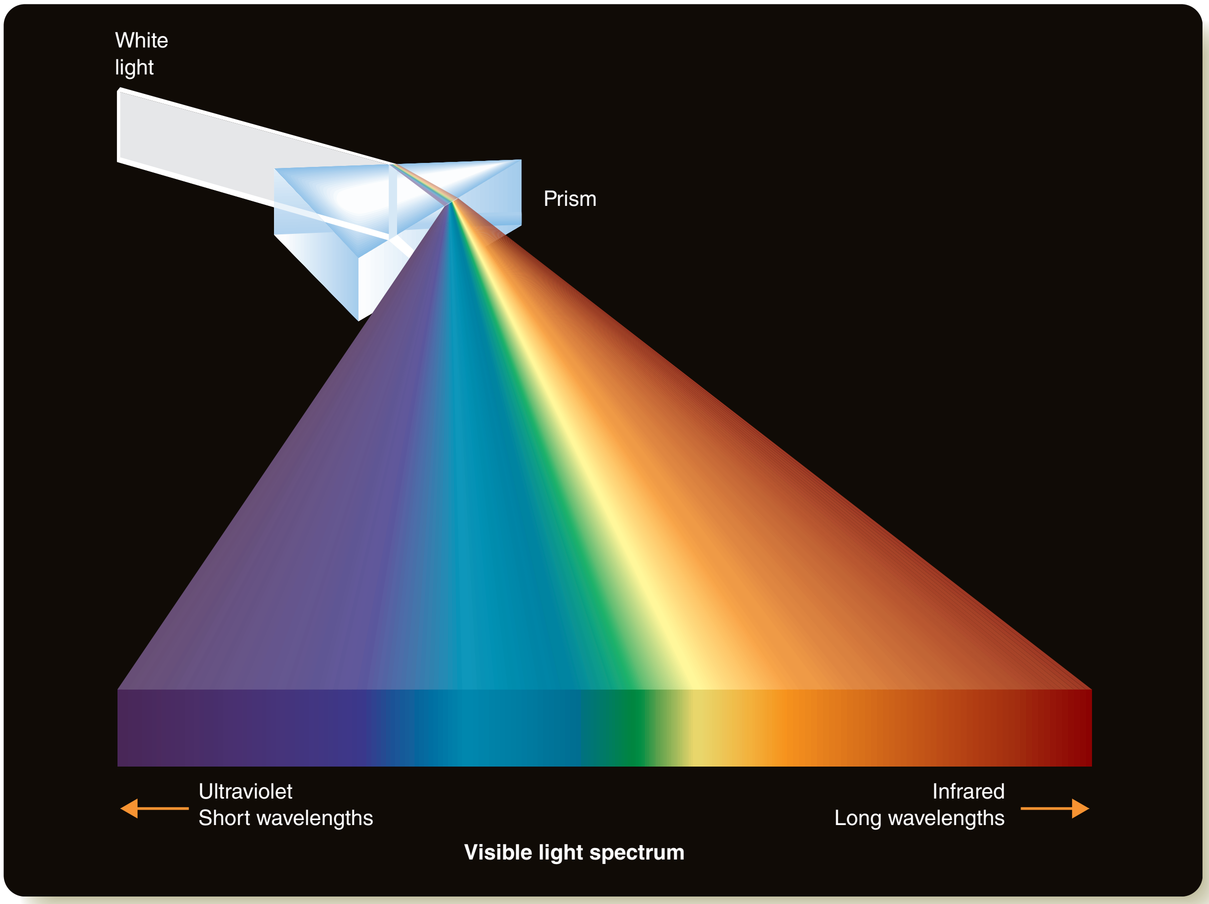


Figure 5-22

A prism bends light by an angle that depends on the wavelength of the light. Short wavelengths bend most and long wavelengths least. Thus, white light passing through a prism is spread into a spectrum.

Adaptive Optics

You have already learned about active optics, a technique to adjust the shape of telescope optics slowly, compensating for effects of changing temperature as well as gravity bending the mirror when the telescope points at different locations in the sky. **Adaptive optics** is a more sophisticated technique that uses high-speed computers to monitor the distortion produced by turbulence in Earth's atmosphere and rapidly alter some optical components to correct the telescope image, sharpening a fuzzy blob into a crisp picture. The resolution of the image is still limited by diffraction in the telescope, but removing much of the seeing distortion produces a dramatic improvement in the detail that is visible (■ Figure 5-23a).

To monitor the distortion in an image, adaptive optics systems must look at a fairly bright star in the field of view, and there isn't always such a star properly located near a target object such as a faint galaxy. In that case, astronomers can point a laser in a direction very close to that of their target object (Figure 5-23b). The laser causes gas in Earth's upper atmosphere to glow, producing an artificial star called a **laser guide star** in the field of view. The adaptive optics system can use information from the changing shape of the artificial star's image to correct the image of the fainter target.

Earlier you learned about huge existing and planned optical telescopes composed of segmented mirrors ten or more meters in diameter. Those telescopes would be much less useful without the addition of adaptive and active optics.

Interferometry

One of the reasons astronomers build big telescopes is to increase resolving power, and astronomers have been able to achieve very high resolution by connecting multiple telescopes together to work, in a sense, as if they comprised a single very large telescope. This method of synthesizing a "virtual" large telescope from

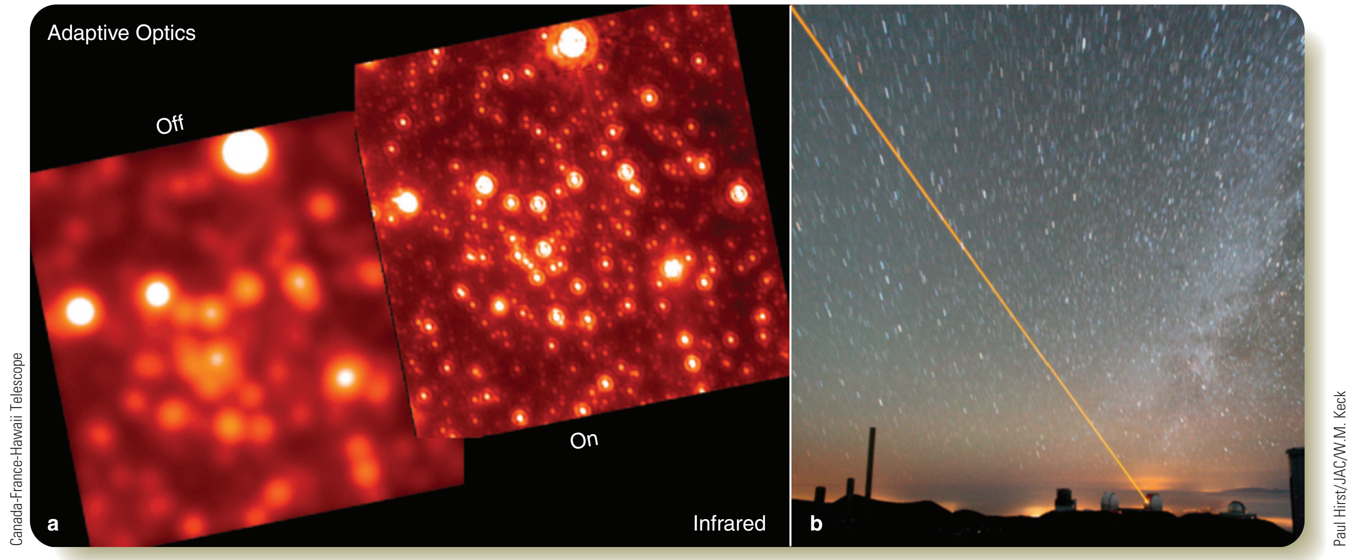


Figure 5-23

(a) In these images of the center of our galaxy, the adaptive optics system was turned "Off" for the left image and "On" for the right image. Not only are the images of stars sharper in the "On" image, but because the light is focused into smaller images, fainter stars are visible. (b) The laser beam shown leaving one of the Keck telescopes produces an artificial star in the field of view, and the adaptive optics system uses that laser guide star as a reference to reduce seeing distortion in the entire image.

two or more smaller telescopes is known as **interferometry** (■ Figure 5-24). The images from such an interferometric telescope are not limited by the diffraction fringes of the individual small telescopes but rather by the diffraction fringes of the much larger virtual telescope.

In an interferometer, light from the separate telescopes must be brought together and combined carefully. The path that each light beam travels must be controlled so that it does not vary more than a small fraction of the light's wavelength. Turbulence in Earth's atmosphere constantly distorts the light, so high-speed computers must continuously adjust the light paths.

As you already know, the resolving power of a radio telescope is relatively low. A dish 30 m in diameter receiving radiation with a wavelength of 21 cm has a resolution of only about 0.5 degrees. In other words, such a radio telescope would be unable to detect any details in the sky smaller than the apparent size of the moon. But, ironically, because long-wavelength radio waves are relative easy to manipulate, radio astronomers were the first to learn how to combine two or more telescopes to form an interferometer capable of much higher resolution than a single telescope.

Radio interferometers can be quite large. The Very Large Array (VLA) consists of 27 dish antennas spread in a Y-shape across the New Mexico desert (■ Figure 5-25a). In combination, they have the resolving power of a radio telescope 36 km (22 mi) in diameter. The VLA can resolve details smaller than 1 arc

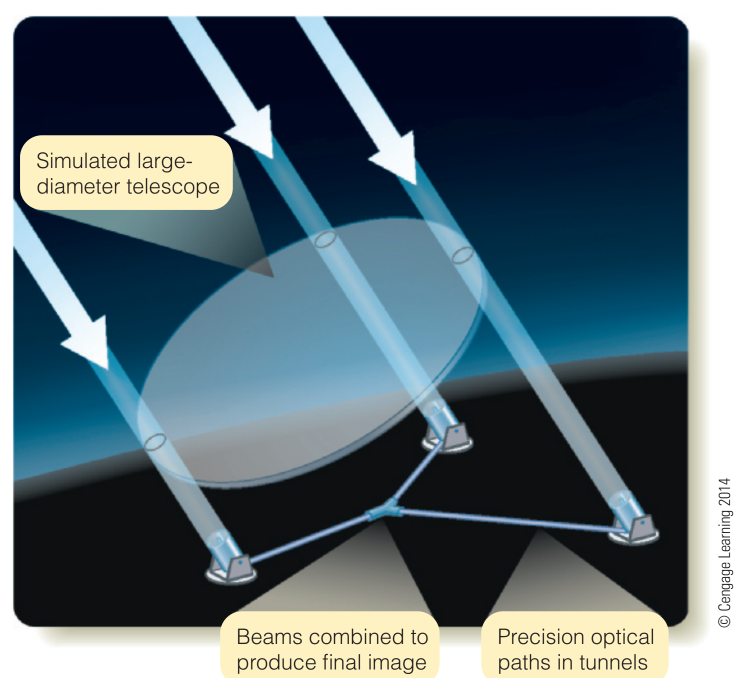


Figure 5-24

In an astronomical interferometer, smaller telescopes can combine their light to simulate a larger telescope with a resolution set by the separation of the smaller telescopes.

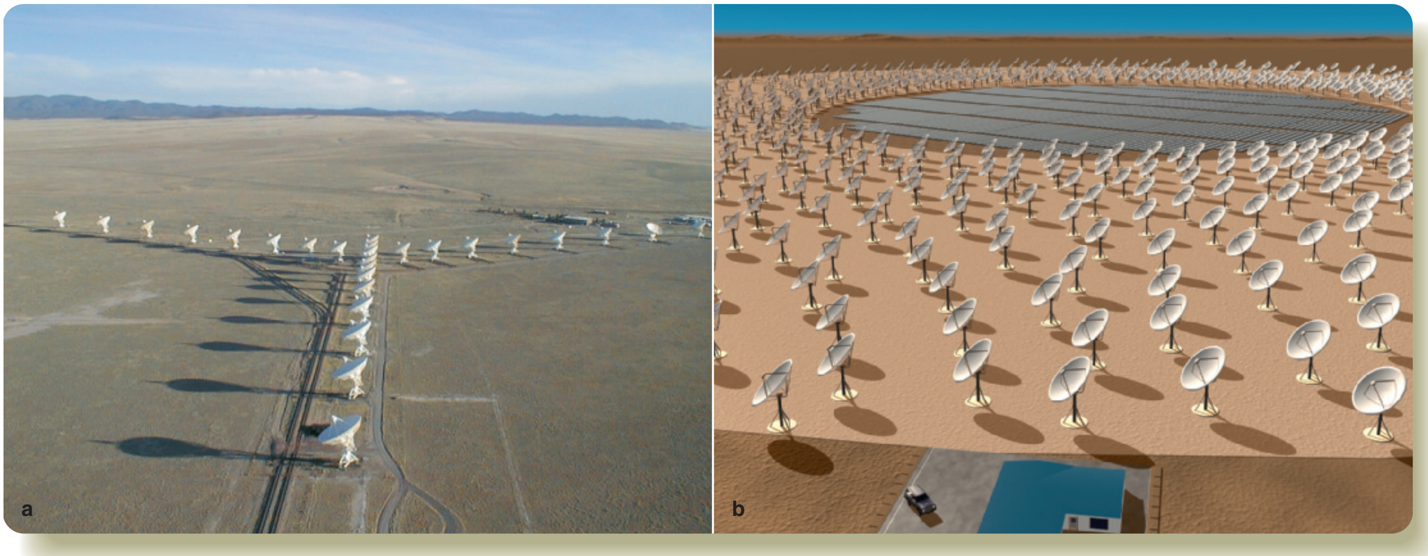


Image courtesy of NRAO/AUI

© Xilostudios/SKA Program Development Office

Figure 5-25

(a) The Very Large Array (VLA) radio telescope uses 27 radio dishes, which can be moved to different positions along a Y-shaped set of tracks across the New Mexico desert. They are shown here in the most compact arrangement. Signals from the dishes are combined to create very-high-resolution radio maps of celestial objects. (b) The proposed Square Kilometer Array (SKA) will have a concentration of radio dishes near its center with more dishes scattered up to 3000 km away.

second, rivaling the performance of a large optical telescope at a good site. Eight new dish antennas being added across New Mexico will give the VLA ten times better resolving power. Another large radio interferometer, the Very Long Baseline Array (VLBA), consists of matched radio dishes spread from Hawai'i to the Virgin Islands and has an effective diameter almost as large as Earth. Radio astronomers are now planning the Square Kilometer Array, which will contain a huge number of radio dishes totaling a square kilometer of collecting area and spread to a diameter of at least 6000 kilometers (Figure 5-25b). These giant radio interferometers depend on state-of-the-art computers to combine signals properly and create radio maps.

Recall that the wavelength of light is very short, roughly 0.0005 mm, so building optical interferometers is one of the most difficult technical problems that astronomers face, but the challenge has been met in several instances. The European Very Large Telescope (VLT) shown in Figure 5-13a consists of four 8.2-m telescopes that can operate separately, but the light they collect, along with light from three 1.8-m telescopes on the same mountaintop, can be brought together through underground tunnels. The resulting optical interferometer, known as the VLTI, can provide the resolution (but, of course, not the light-gathering power) of a telescope 200 m (660 ft, more than two football fields) in diameter. Astronomers using the VLTI in 2009 made an image of the red giant star T Leporis with a resolution of 0.004 arc seconds, equivalent to being able to discern a two-story house on the moon.

Other facilities, such as the CHARA (Center for High Angular Resolution Astronomy) telescope array on Mount Wilson, the two Keck 10-m telescopes in Hawai'i, and the Large Binocular Telescope in Arizona, also can work as interferometers.

Although turbulence in Earth's atmosphere can be partially averaged out in an interferometer, plans are being made to put interferometers in space to avoid atmospheric turbulence altogether. Designs have been proposed in the United States for the *Space Interferometry Mission (SIM)* and in Europe for *Darwin*, both of which would be multispace telescope interferometers capable of studying everything from the cores of active galaxies to planets orbiting nearby stars with extremely high angular resolution. Those projects will not be started until the 2020s at the earliest.

Nonelectromagnetic Astronomy

This chapter is focused on analyzing electromagnetic radiation from space. Other types of energy also arrive here bearing information from the rest of the universe and deserve a brief mention.

Cosmic rays are subatomic particles traveling through space at tremendous velocities. Almost no cosmic rays reach the ground, but some of them smash into gas atoms in Earth's upper atmosphere, and fragments of those atoms shower down to the ground. Those secondary cosmic rays are passing through you as you read this sentence. Other types of particles from space interact weakly

and seldom with Earth atoms, so huge detectors must be built to catch and count them. Detectors for some kinds of cosmic rays have been carried on balloons or launched into orbit, while others have been built deep underground. Astronomers are not yet sure what produces cosmic rays because their original sources are difficult to locate. There is evidence that at least a few high-energy cosmic rays are produced by violent explosions of dying stars or supermassive black holes at the centers of galaxies. You will meet these exotic objects again in future chapters.

Einstein's general theory of relativity (see Chapter 16) predicts that another type of nonelectromagnetic signal called gravity waves should be produced by any mass in the universe that accelerates. The existence of gravity waves has been inferred from some observations but they have never been detected directly because they are extremely weak. The LIGO (Laser Interferometer Gravitational Wave Observatory) is a ground-based facility intended to be sensitive enough to detect cosmic gravity waves after an advanced version is completed in 2013. *LISA (Laser Interferometry Space Antenna)* is its planned highly sensitive space-based counterpart.

What Are We? Curious

Telescopes are creations of curiosity. You look through a telescope to see more and to understand more. The unaided eye is a detector with limited sensitivity, and the history of astronomy is the history of bigger and better telescopes gathering more and more light to search for fainter and more distant objects.

The old saying "Curiosity killed the cat" is an insult to the cat and to curiosity. We humans are curious, and curiosity is a noble trait—the mark of an active, inquiring mind. At the limits of human curiosity lies the fundamental question, "What are we?" Telescopes extend and amplify our senses, but they also allow us to extend and amplify our curiosity about the universe around us.

When people find out how something works, they say their curiosity is satisfied. Curiosity is an appetite like hunger or thirst, but it is an appetite for understanding. As astronomy expands our horizons and we learn about how distant stars and galaxies form and evolve, we feel satisfaction partly because we are learning about ourselves and our place in the universe. We are beginning to understand what we are.

Study and Review

Summary

- Light is the visible form of **electromagnetic radiation (p. 75)**, an electric and magnetic disturbance that transports energy at the speed of light. The **wavelength (p. 75)** of visible light is usually measured in **nanometers (p. 76)** (10^{-9} m) or **angstroms (p. 76)** (10^{-10} m), and ranges from 400 nm to 700 nm (4000 to 7000 Å).
- **Frequency (p. 75)** is the number of waves that pass a stationary point in 1 second. The frequency of an electromagnetic wave equals the speed of light divided by the wave's wavelength.
- A **photon (p. 76)** is a packet of light waves that can act as a particle or as a wave. The energy carried by a photon is proportional to its frequency and inversely proportional to its wavelength.
- A **spectrum (p. 76)** is a display of light sorted and viewed or recorded in order of wavelength. The complete electromagnetic spectrum includes **gamma rays (p. 76)**, **X-rays (p. 76)**, **ultraviolet (UV) (p. 76)** radiation, visible light, **infrared (IR) (p. 76)** radiation, **microwaves (p. 76)**, and **radio waves (p. 76)**.
- Gamma-ray, X-ray, and ultraviolet radiation have shorter wavelengths, higher frequency, and carry more energy per photon than visible light. Infrared rays, microwaves, and radio waves have longer wavelengths and lower frequency and carry less energy per photon than visible light.
- Earth's atmosphere is fully transparent in only two **atmospheric windows (p. 77):** visible light and radio.
- **Refracting telescopes (p. 78)** use a **primary lens (p. 78)** to bend the light and focus it into an image. **Reflecting telescopes (p. 78)** use a **primary mirror (p. 78)** to focus the light. The image produced by the telescope's primary lens or mirror can be magnified by an **eyepiece (p. 78).** Lenses and mirrors with short **focal lengths (p. 78)** must be strongly curved and are more expensive to grind to an accurate shape.
- Because of **chromatic aberration (p. 78),** refracting telescopes cannot bring all colors to the same focus, resulting in color fringes around the images. An **achromatic lens (p. 78)** partially corrects for this, but such lenses are expensive and cannot be made much larger than about 1 m in diameter.
- Reflecting telescopes are easier to build and less expensive than refracting telescopes of the same diameter. Also, reflecting telescopes do not suffer from chromatic aberration. Most large **optical telescopes (p. 79)** and all **radio telescopes (p. 79)** are reflecting telescopes.
- **Light-gathering power (p. 79)** refers to the ability of a telescope to produce bright images. **Resolving power (p. 80)** refers to the ability of a telescope to resolve fine detail. **Diffraction fringes (p. 80)** in an image, caused by the interaction of light waves with the telescope's apertures, limit the detail visible. **Magnifying power (p. 82),** the ability to make an object look bigger, is a less important telescope power because it is not a property of the telescope itself but can be changed simply by changing the eyepiece.
- Astronomers build optical observatories on remote, high mountains for two reasons. (1) Turbulence in Earth's atmosphere blurs the image of an astronomical object, a phenomenon that astronomers refer to as **seeing (p. 80).** Atop a mountain, the air is relatively steady, and the seeing is better. (2) Observatories are located far from cities to avoid **light pollution (p. 83).** Radio telescopes are also located far from cities to avoid human-produced radio noise.
- In a reflecting telescopes, light first comes to a focus at the **prime focus (p. 88)**, but **secondary mirrors (p. 88)** can direct light to other locations such as the **Cassegrain focus (p. 88).** The **Newtonian focus (p. 88)** and **Schmidt-Cassegrain focus (p. 88)** locations are used in some telescopes.
- Because Earth rotates, telescopes must have a **sidereal drive (p. 89)** to remain pointed at celestial objects. An **equatorial mount (p. 89)** with a **polar axis (p. 89)** is the simplest way to accomplish this. An **alt-azimuth mount (p. 89)** can support a more massive telescope but requires computer control to compensate for Earth's rotation.
- Very large telescopes can be built with **active optics (p. 89)** to control the shapes of mirrors that are thin and flexible or composed of segments. Such thin or segmented mirrors weigh less, are easier to support, and cool faster at nightfall, but their shapes need to be adjusted gradually and continuously to maintain a good focus.
- Earth's atmosphere distorts and blurs images. Telescopes in orbit are above this seeing distortion and are limited only by diffraction in their optics. Earth's atmosphere absorbs gamma rays, X-rays, ultraviolet, far-infrared, and microwave light. To observe at these wavelengths, telescopes must be located at high altitudes or in space.
- Astronomers in the past used **photographic plates (p. 93)** to record images at the telescope, and **photometers (p. 93)** to precisely measure the brightness of celestial objects. Modern electronic systems such as **charge-coupled devices (CCD) (p. 93)** and other types of **array detectors (p. 93)** have replaced both photographic plates and photometers in most applications.
- Electronic detectors have the advantage that data from them is automatically **digitized (p. 93)** in numerical format and can be easily recorded and manipulated. Astronomical images in digital form can be computer-enhanced to produce **representational-color** images (also called **false-color** images) **(p. 93)** that bring out subtle details.
- **Spectrographs (p. 94)** using prisms or a **grating (p. 94)** spread starlight out according to wavelength to form a spectrum revealing hundreds of **spectral lines (p. 94)** produced by atoms and molecules in the object being studied.
- **Adaptive optics (p. 95)** techniques involve measuring seeing distortions caused by turbulence in Earth's atmosphere and partially canceling out those distortions by rapidly altering some of the telescope's optical components. In some facilities a powerful laser beam is used to produce an artificial **laser guide star (p. 95)** high in Earth's atmosphere that can be monitored by an adaptive optics system.
- **Interferometry (p. 96)** refers to the technique of connecting two or more separate telescopes to act as a single large telescope that has a resolution equivalent to that of a telescope as large in diameter as the separation between the individual telescopes. The first working interferometers were composed of multiple radio telescopes.
- **Cosmic rays (p. 97)** are not electromagnetic radiation; they are subatomic particles such as electrons and protons traveling at nearly the speed of light, arriving from mostly unknown cosmic sources.

Review Questions

1. Why would you not plot sound waves in the electromagnetic spectrum?
2. If you had limited funds to build a large telescope, which type would you choose, a refractor or a reflector? Why?
3. Why do nocturnal animals usually have large pupils in their eyes? How is that related to the way astronomical telescopes work?
4. Why do optical astronomers often put their telescopes at the tops of mountains, while radio astronomers sometimes put their telescopes in deep valleys?
5. What are the advantages of making a telescope mirror thin? What problems does that cause?
6. Small telescopes are often advertised as "200 power" or "magnifies 200 times." How would you improve such advertisements?
7. Why do radio telescopes have relatively poor resolving power?
8. The moon has no atmosphere at all. What advantages would you have if you built an observatory on the lunar surface?
9. Why must telescopes observing at far-infrared wavelengths be cooled to low temperatures?
10. What purpose do the colors in a representational-color image or map serve?
11. What might you detect with an X-ray telescope that you could not detect with an infrared telescope?
12. How is the phenomenon of chromatic aberration related to how a prism spectrograph works?
13. Why would radio astronomers build identical radio telescopes in many different places around the world?
14. **How Do We Know?** How is the resolution of an astronomical image related to the precision of a measurement?

Discussion Questions

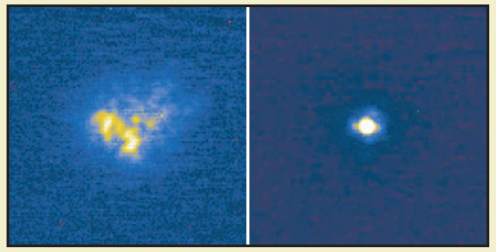
1. Why does the wavelength response of the human eye match the visual window of Earth's atmosphere so well?
2. Most people like beautiful sunsets with brightly glowing clouds, bright moonlit nights, and twinkling stars. Astronomers don't. Why?

Problems

1. The thickness of the plastic in plastic bags is about 0.001 mm. How many wavelengths of red light is that?
2. What is the wavelength of radio waves transmitted by a radio station with a frequency of 100 million cycles per second?
3. Compare the light-gathering power of one of the 10-m Keck telescopes with that of a 0.5-m telescope.
4. How does the light-gathering power of one of the Keck telescopes compare with that of the human eye? Assume that the pupil of your eye can open to a diameter of about 0.8 cm in dark conditions.
5. What is the resolving power of a 25-cm (10-in.) telescope at a wavelength of 550 nm (in the middle of the visual band)? What do two stars 1.5 arc seconds apart look like through this telescope at that wavelength?
6. Most of Galileo's telescopes were only about 2 cm in diameter. Should he have been able to resolve the two stars mentioned in Problem 5?
7. How does the resolving power of the 5-m telescope on Mount Palomar near San Diego compare with that of the 2.5-m *Hubble Space Telescope*? Why does the HST generally still outperform the Palomar 5-m telescope?
8. If you build a telescope with a focal length of 1.3 m, what focal length does the eyepiece need to give a magnification of 100 times?
9. Astronauts observing from a space station need a telescope with a light-gathering power 15,000 times that of the dark-adapted human eye (*Note:* see problem 4), capable of resolving detail as small as 0.1 arc seconds at a wavelength of 550 nm, and having a magnifying power of 250. Design a telescope to meet their needs. Could you test your design by using it to observe stars from the surface of Earth?
10. A spy satellite orbiting 400 km above Earth is supposedly capable of counting individual people in a crowd in visual-wavelength images. Assume that the middle of the visual wavelength band is a wavelength of 550 nm. Assume an average person seen from above has a size of 0.7 meters. Use the small-angle formula (Chapter 3) and the formula for telescope resolving power to estimate the minimum diameter of the telescope the satellite must carry.

Learning to Look

1. The two images below show a star before and after an adaptive optics system attached to the telescope was switched on. What causes the distortion in the first image, and how do adaptive optics correct the image?

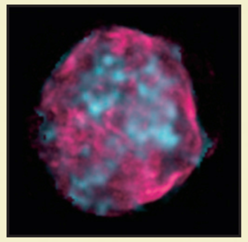


ESO

2. The star images in the photo below are tiny disks, but the diameter of these disks is not related to the diameter of the stars. Explain why the telescope can't resolve the diameter of the stars. What does cause the apparent diameters of the stars?

NASA, ESA, and G. Meylan (Ecole Polytechnique Federale de Lausanne)

3. The X-ray image below shows the remains of an exploded star. Explain why images recorded by telescopes in space are often displayed in representational ("false") color rather than in the "colors" (wavelengths) received by the telescope.



NASA, CXC, PSU, S. Park

Great Debates

1. *Is Your Countertop Radiating?* The U.S. Environmental Protection Agency has posted information on their website about radiation emitted from granite (www.epa.gov/radiation/tenorm/granite-countertops.html). You can also do an Internet search for "EPA granite radiation" and follow the links. Uranium 238 decays into radon 222 over time. Because of this a radioactive gas may be leaking from granite countertops, and this radiation may give you lung cancer from breathing in the gas over time. Should the public remove their granite or sandstone countertops from their homes even if the EPA concludes that the radiation dose received is small? If the EPA concludes that countertops should be removed, should the government (that is, your tax dollars) pay for the removal? Should the companies who sold the countertops recall all countertops? Should radon kits be issued to all homes containing these countertops regardless of the report's conclusions? Because radon can be eliminated in most homes by opening doors and windows, should the government issue safety pamphlets that contain this information?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

2. *Self-Paying Telescopes.* Installation of large, ground-based telescopes is very expensive, and some governments have combined resources to jointly fund projects. Ultimately the taxpayer pays for these telescopes. If a large ground-based telescope can be built as a solar collector during the day and telescope during the night, it might pay for itself and make a profit over time. Should these telescopes cease to be jointly funded?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

3. *Cosmic Rays.* Cosmic rays are subatomic particles traveling through space at a good fraction of the speed of light. Electrons, protons or hydrogen nuclei, and alpha particles or helium nuclei are examples of these subatomic particles. Cosmic rays are not electromagnetic radiation, although they were once thought to be. Historically, astronomical names are not changed once they are assigned (for example, the Big Bang was neither big nor a bang; black holes are neither black nor holes; gravity waves are not EM radiation, and so on). Should the cosmic ray name be changed to better reflect the physics? Is the name confusing?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

4. *Are Ghosts Real?* Ghosts are thought to be the spirits of dead people. Ghosts are often reported as being heard or "seen" in pictures taken by cameras. Are the ghosts seen in images taken by cameras real objects? How can a ghost be real if it is nonmaterial? How can it interact with EM radiation?
   a. In your argument, state the wavelength bands in Figure 5-2 that human eyes see in and the wavelength bands that cameras see in. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand including images of ghosts in pictures.
   c. Cite your sources.

5. *Electronic Billboards.* A proposed bill on regulation of electronic billboard locations will soon be up for a vote. If an electronic billboard is located near an observatory, then the lights will likely interfere with the data collected by instruments attached to the telescopes. Electronic billboards on roads along the ocean may confuse the navigation of baby turtles, causing them to enter traffic instead of following the moonlight to the ocean. However, lobbyists argue that states can benefit from the increased revenue provided by the billboards. Should locations of electronic billboards be regulated by the federal government?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Refractors and Reflectors
- Resolution and Telescopes

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 2: Light and Matter

Astronomers need to be experts in light. The stars and galaxies are far beyond reach, so astronomers have to make measurements by analyzing the light their telescopes gather. Astronomers have to know how light behaves and how it is emitted by hot matter. Of course, the electromagnetic spectrum includes more types of light than just the type that human eyes can see. The mysteries of the universe are so profound that astronomers must use every wavelength possible to understand such processes as how stars evolve and how planets form.

This lab will introduce you to light waves and give you a chance to do some simple calculations. You will study two of the basic laws of electromagnetic radiation, known as Planck's law and the Wien's law, and you will study how these are related to the luminosities of stars. You will also get a chance to see how light waves interact with atoms. The first step is to understand waves.

You experience several types of waves in your life, some every day. A dramatic kind of wave is a wave in liquid. Imagine yourself in a very large body of water—the ocean or one of the Great Lakes—treading water out beyond where the waves break. The wave has a wavelength, meaning the distance from crest to crest or trough to trough. Let's imagine together that the wavelength of the waves you are floating in is 50 feet. The wave also has a frequency, meaning how many wave crests pass a stationary location per second. The normal units used are waves, or cycles, per second, known as "hertz" (abbreviated Hz), named after the scientist who discovered radio waves. Your water wave might have a crest pass only once every 10 seconds, so its frequency is less than one, that is, 1/10 of a wave passes per second, or 0.1 Hz.

Your wave also has a speed. In this chapter you learned the simple formula that wavelength times frequency equals speed. The units work out: Length times "per second" equals a speed, length per second. The water wave we have constructed in our minds' eyes must have a speed equal to (50 feet) × (1/10 per second), which works out as 5 feet per second.

Note that your water wave moves horizontally, but you don't. You bob up and down, as does anything else floating next to you in the water such as a stick of wood. If you could see individual water molecules you would know the same is true about them as well. The wave is a pattern, carrying energy, perhaps from a distant storm, and the pattern moves horizontally, but it causes the matter it encounters, including you, to oscillate vertically, at right angles to the wave pattern's motion. This is called a "transverse" wave. You are imagining experiencing these waves out beyond "the break," the point at which the depth of the water is less than the wavelength of the waves coming toward the shore. There, the behavior of the wave changes. The wave height increases, and the water actually begins to have bulk motion horizontally, to the delight of surfers.

Light also travels as a wave. But what is "waving"? You have learned that light, and electromagnetic radiation in general, is an oscillating disturbance of electrical and magnetic fields. Light is a transverse wave, like water waves. Light shining on matter makes charged particles such as electrons wiggle back and forth at right angles to the direction the wave pattern is moving. Sound also travels as a wave, a series of compressions and decompressions of the air. In this case the oscillating motion of the matter (air molecules) caused by the traveling wave pattern is parallel to the wave's motion. That is a different type of wave, called a "longitudinal" wave. Nevertheless, sound waves and light waves both follow the rule that wavelength times frequency equals speed. Of course the speeds of water waves, sound waves, and light waves are very different.

Finally, a wave has an amplitude, meaning its strength, which for a water wave is usually expressed as the vertical distance between the bottom of the trough and the top of the crests. The amplitude of a sound wave is experienced as its loudness, and the amplitude of a light wave is its brightness.

Section 1 of Virtual Astronomy Lab 2, "Light and Matter," continues your introduction to electromagnetic radiation begun in this chapter. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

6

Atoms and Spectra

Guidepost

In the previous chapter you learned how telescopes gather light, cameras record images, and spectrographs spread light into spectra (plural of *spectrum*). Now you are ready to understand why astronomers make great efforts to gather and analyze spectra of the objects they study. Here you will find answers to three important questions:

- ► **How do atoms interact with light to produce spectra?**
- ► **What are the types of spectra you can observe?**
- ► **What can you learn from spectra of celestial objects?**

This chapter marks a change in the way you will look at nature. Up to this point, you have been considering what you can see with your eyes alone or aided by telescopes. In this chapter, you will begin using modern **astrophysics,** linking astronomical observations to physics observations and theories, to investigate secrets that lie beyond what you can see directly.

In the chapters that follow you will study stars, galaxies, and planets, often using the rich information derived from their spectra.

A display of the aurora borealis ("northern lights"). Gas in Earth's upper atmosphere is excited by electrical currents caused by charged particles in the solar wind interacting with Earth's magnetic field. The auroral glow shows spectral emission lines of ionized oxygen (green) and nitrogen (red).

© Anna Henley/Roam Images/Jupiterimages

Awake! for Morning in the Bowl of Night
Has flung the Stone that puts the Stars to Flight:
And Lo! the Hunter of the East has caught
The Sultan's Turret in a Noose of Light.

THE RUBÁIYÁT OF OMAR KHAYYÁM,
TRANSLATION BY EDWARD FITZGERALD

THE UNIVERSE IS populated with brilliant stars illuminating exotic planets and fabulously beautiful clouds of glowing gas. But beyond the members of our tiny local solar system, they are all out of reach for the foreseeable future. No human space probe has visited another star, and no telescope can directly examine the insides of any celestial object. The information you can obtain about most of the universe is contained in the light reaching you across space (■ Figure 6-1).

NASA, Andrew Fruchter and the ERO Team (Sylvia Baggett (STScI), Richard Hook (ST-ECF), Zoltan Levay (STScI))

■ Figure 6-1

What's going on here? The sky is filled with beautiful and mysterious objects that lie far beyond your reach—in the case of the nebula NGC 2392, also called the Eskimo Nebula, about 5000 ly beyond your reach. The only way to understand such objects is by analyzing their light. Such an analysis reveals that this object is a dying star surrounded by the expanding shell of gas it ejected a few thousand years ago. You will learn more about this phenomenon in a later chapter.

Earthbound humans knew almost nothing about the composition of celestial objects until the early 19th century. First, the German optician Joseph von Fraunhofer studied the spectrum of the sun and discovered that it is interrupted by more than 600 dark lines representing colors missing from the sunlight Earth receives. Then, other scientists performed laboratory experiments showing that those spectral lines are related to the presence of various atoms in the sun's atmosphere. Finally, astronomers observed that the spectra of other stars have similar patterns of lines, which opened a window to real understanding of how the sun and stars are related. In this chapter you will look through that window, seeing how the sun and other stars produce light and how atoms in the atmospheres of stars, planets, and gas clouds in space interact with light to cause spectral lines. Once you understand that, you will know how astronomers determine the chemical composition of distant objects, as well as measure motions of gas in and around them.

6-1 Atoms

ATOMS IN STARS AND PLANETS leave their fingerprints on the light they emit. By first reviewing what atoms are and then learning how they interact with light, you can understand how the spectra of objects in space are decoded.

A Model Atom

To think about how atoms interact with light, you need a working model of an atom. In Chapter 2, you used a model of the sky, the celestial sphere. In this chapter, you will begin your study of atoms by creating a mental model of an atom. Remember that such a model can have practical value without being true. The stars are not actually attached to a sphere surrounding Earth, but to navigate a ship or point a telescope, it is useful to pretend they are. The electrons in an atom are not actually little beads orbiting the nucleus the way planets orbit the sun, but for some purposes it is useful to picture them as such.

Your model atom contains a positively charged **nucleus** at the center, which consists of two kinds of particles: **Protons** carry a positive electrical charge, and **neutrons** have no charge. That means the nucleus has a net positive charge. The nucleus of the model atom is surrounded by a cloud of **electrons** that are low-mass particles with negative charges. Normally the number of electrons equals the number of protons, and the positive and negative charges balance to produce a neutral atom.

A single atom is not a massive object. A hydrogen atom, for example, has a mass of only 1.67×10^{-27} kg, about a trillionth of a trillionth of a gram. Protons and neutrons have masses

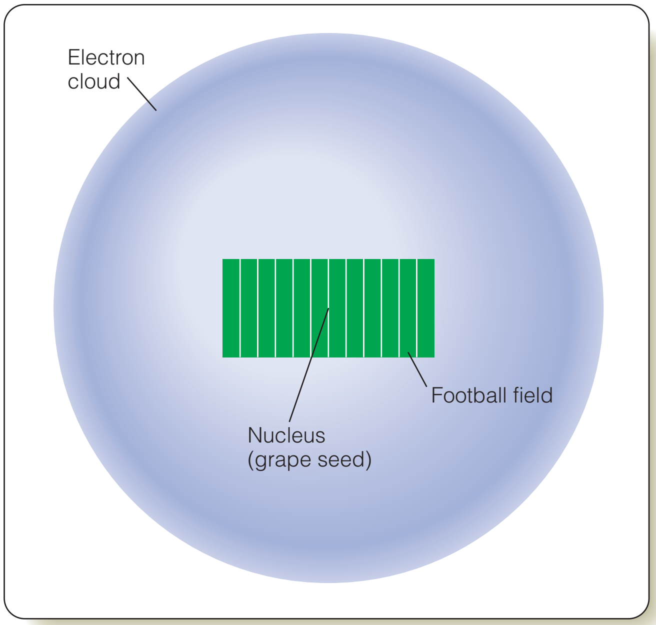


Figure 6-2

Magnifying a hydrogen atom by 10^{12} makes the nucleus the size of a grape seed and the diameter of the electron cloud almost 3 times larger than an American football field.

almost 2000 times greater than that of an electron, so most of the mass of an atom lies in the nucleus.

An atom is mostly empty space. To see this, imagine constructing a simple scale model of a hydrogen atom. Its nucleus is a single proton with a diameter of approximately 0.0000016 nm, or 1.6×10^{-15} m. If you multiply this by one trillion (10^{12}), you can represent the nucleus of your model atom with something about 0.16 cm in diameter—a grape seed would do. The region of a hydrogen atom that contains the electron has a diameter of about 0.24 nm, or 2.4×10^{-10} m. Multiplying by a trillion increases the diameter to about 240 m, or almost 3 football fields laid end to end (■ Figure 6-2). When you imagine a grape seed in the middle of a sphere three football fields in diameter, you can see that an atom is mostly empty space.

Now you can consider a **Common Misconception.** Most people, without thinking about it much, imagine that matter is solid, but you have seen that atoms are mostly empty space. The chair you sit on, the floor you walk on, are mostly not there. A later chapter will reveal to you what happens to a star when most of the empty space gets squeezed out of its atoms.

Different Kinds of Atoms

There are over a hundred chemical elements. The number of protons in the nucleus of an atom determines which element it is. For example, a carbon atom has six protons in its nucleus. An atom with one more proton than that is nitrogen, and an atom with one fewer proton is boron.

Although an atom of a given element always has the same number of protons in its nucleus, the number of neutrons is less restricted. For instance, if you added a neutron to a carbon nucleus, it would still be carbon, but it would be slightly heavier. Atoms that have the same number of protons but a different number of neutrons are **isotopes**. Carbon has two stable isotopes. One contains six protons and six neutrons for a total of 12 particles and is thus called carbon-12. Carbon-13 has six protons and seven neutrons in its nucleus.

The number of electrons in an atom of a given element can vary. Protons and neutrons are bound tightly into the nucleus, but the electrons are held loosely in the electron cloud. Running a comb through your hair creates a static charge by removing a few electrons from their atoms. An atom that has lost or gained one or more electrons is said to be **ionized** and is called an **ion.** A neutral carbon atom has six electrons that balance the positive charge of the six protons in its nucleus. If you ionize the atom by removing one or more electrons, the atom is left with a net positive charge. Under other circumstances, an atom may capture one or more extra electrons, giving it more negative charges than positive. Such a negatively charged atom is also considered an ion.

Atoms that collide may form bonds with each other by exchanging or sharing electrons. Two or more atoms bonded together form a **molecule**. Atoms do collide in stars, but the high temperatures cause violent collisions that are unfavorable for chemical bonding. Only in the coolest stars are the collisions gentle enough to permit the formation of chemical bonds. The presence of molecules such as titanium oxide (TiO) detected in some stars is a clue that those stars are very cool compared with other stars. In later chapters, you will see that molecules also can form in cool gas clouds in space and in the atmospheres of planets.

Electron Orbits

So far you have been considering the cloud of electrons in atoms only in a general way. Now it is time to be more specific about how electrons behave within the cloud on the way to understanding how light interacts with atoms.

Electrons are bound to the atom by the attraction between their negative charge and the positive charge on the nucleus. This attraction is known as the **Coulomb force**, after the French physicist Charles-Augustin de Coulomb (1736–1806). To ionize an atom, you need a certain amount of energy to pull an electron completely away from the nucleus. This energy is the electron's **binding energy**, the energy that holds it to the atom.

How Do We Know? | 6-1

Quantum Mechanics

How can you understand nature if it depends on the atomic world you cannot see? You can see objects such as stars, planets, aircraft carriers, and hummingbirds, but you can't see individual atoms. As scientists apply the principle of cause and effect, they study the natural effects they can see and work backward to find the causes. Invariably that quest for causes leads back to the invisible world of atoms.

Quantum mechanics is the set of rules that describe how atoms and subatomic particles behave. On the atomic scale, particles behave in ways that seem unfamiliar. One of the principles of quantum mechanics specifies that you cannot know simultaneously the exact location and motion of a particle. This is why physicists prefer to go one step beyond the simple atomic model that imagines electrons as particles following orbits, and instead describe the electrons in an atom as if they each were a cloud of negative charge. That's a better model.

This raises some serious questions about reality. Is an electron really a particle at all? If you can't know simultaneously the position and motion of a specific particle, how can you know how it will react to a collision with a photon or another particle? The answer is that you can't know, and that seems to violate the principle of cause and effect.

Many of the phenomena you can see depend on the behavior of huge numbers of atoms, and quantum mechanical uncertainties average out. Nevertheless, the ultimate causes that scientists seek lie at the level of atoms, and modern physicists are trying to understand the nature of the particles that make up atoms. That is one of the most exciting frontiers of science.

©Antonio V. Oquias/Shutterstock.com

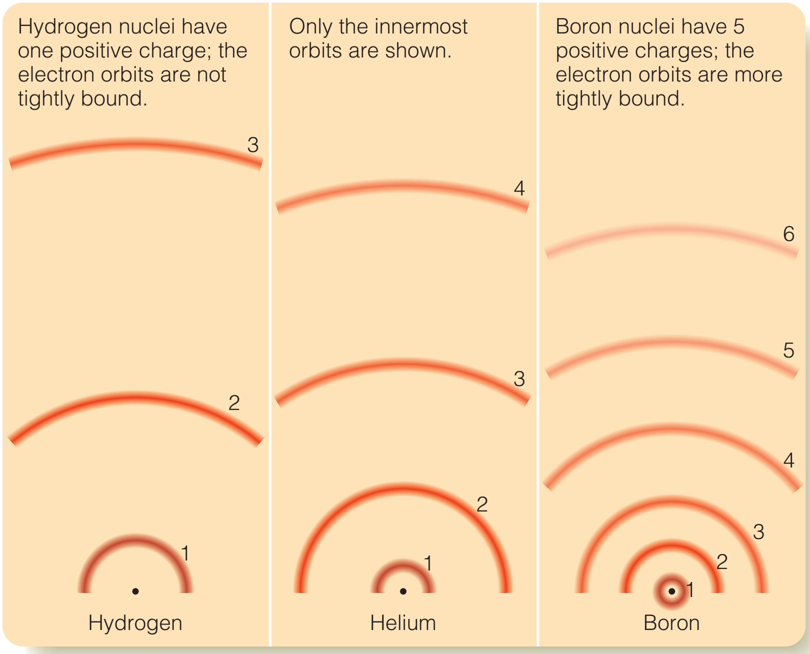
The world you see, including these neon signs, is animated by the properties of atoms and subatomic particles.

The size of an electron's orbit is related to the energy that binds it to the atom. If an electron orbits close to the nucleus, it is tightly bound, and a large amount of energy is needed to pull it away. In other words, its binding energy is large. An electron orbiting farther from the nucleus is held more loosely, and less energy is needed to pull it away. That means it has small binding energy.

Nature permits atoms only certain amounts (quanta) of binding energy. The laws that describe how atoms behave are called the laws of **quantum mechanics** (**How Do We Know? 6-1**). Much of this discussion of atoms is based on the laws of quantum mechanics that were discovered by physicists early in the 20th century.

Because atoms can have only certain amounts of binding energy, your model atom can have orbits only of certain sizes, called **permitted orbits**. These are like steps in a staircase: You can stand on the number-one step or the number-two step, but not on the number-one-and-one-quarter step—there isn't one. The electron can occupy any permitted orbit, but there are no orbits in between.

The arrangement of permitted orbits depends primarily on the charge of the nucleus, which in turn depends on the number of protons. Consequently, each chemical element—each type of atom—has its own pattern of permitted orbits (■ Figure 6-3). Isotopes of the same elements have nearly the same pattern because they have the same number of protons in their nuclei but slightly different masses. Ionized atoms, with altered electrical charges, have orbital patterns that differ greatly from their un-ionized forms.



© Cengage Learning 2014

■ **Figure 6-3**

The electron in an atom may occupy only certain permitted orbits. Because different elements have different charges on their nuclei, the elements have different, unique patterns of permitted orbits.

SCIENTIFIC ARGUMENT

How many hydrogen atoms would it take to cross the head of a pin?

This is not a frivolous question. In answering it, you will discover how small atoms really are, and you will see how powerful physics and mathematics can be as a way to understand nature. Many scientific arguments are convincing because they have the precision of mathematics. To begin, assume that the head of a pin is about 1 mm in diameter—that is, 0.001 m. The size of a hydrogen atom is represented by the diameter of the electron cloud, roughly 0.24 nm. Because 1 nm equals 10^{-9} m, you can multiply and discover that 0.24 nm equals 2.4×10^{-10} m. To find out how many atoms would stretch 0.001 m, you can divide the diameter of the pinhead by the diameter of an atom. That is, divide 0.001 m by 2.4×10^{-10} m, and you get 4.2×10^6. It would take 4.2 million hydrogen atoms lined up side by side to cross the head of a pin.

Now you can see how tiny an atom is and also how powerful a bit of physics and mathematics can be. It reveals a view of nature beyond the capability of your eyes. Now build an argument using another bit of arithmetic. **How many hydrogen atoms would you need to add up to the mass of a paper clip (1 g)?**

6-2 Interactions of Light and Matter

If light and matter did not interact, you would not be able to see these words. In fact, you would not exist, because, among other problems, photosynthesis would be impossible, so there would be no grass, wheat, bread, beef, cheeseburgers, or any other kind of food. The interaction of light and matter makes life possible, and it also makes it possible for you to understand the universe.

You have already been considering a model hydrogen atom. Now you can use that model as you begin your study of light and matter. Hydrogen is both simple and common: Roughly 90 percent of all atoms in the universe are hydrogen.

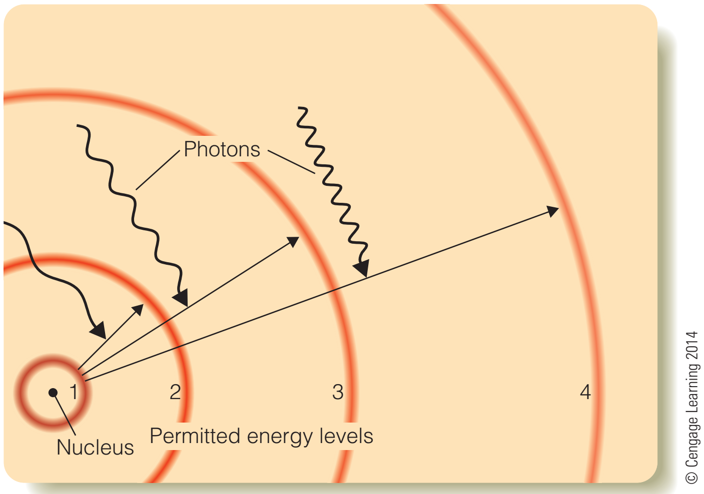
The Excitation of Atoms

Each electron orbit in an atom represents a specific amount of binding energy, so physicists commonly refer to the orbits as **energy levels**. Using this terminology, you can say that an electron in its smallest and most tightly bound orbit is in its lowest permitted energy level, which is called the atom's **ground state**. You could move the electron from one energy level to another by supplying enough energy to make up the difference between the two energy levels. It would be like moving a package from a low shelf to a high shelf; the greater the distance between the shelves, the more energy you would need to raise the package. The amount of energy needed to move the electron is the energy difference between the two energy levels.

If you move an electron from a low energy level to a higher energy level, the atom becomes an **excited atom**. That is, you have added energy to the atom by moving its electron outward from the nucleus. One way an atom can become excited is by collision. If two atoms collide, one or both may have electrons knocked into a higher energy level. This happens very commonly in hot gas, where atoms move rapidly and collide often.

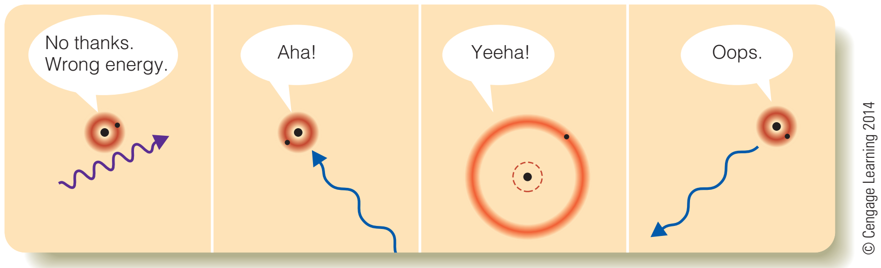
Another way an atom can become excited is to absorb a photon. As you learned in the previous chapter, a photon is a bundle of electromagnetic waves with a specific energy. Only a photon with exactly the right amount of energy can move the electron from one level to another. If the photon has too much or too little energy, that atom cannot absorb it. Because the energy of a photon depends on its wavelength, only photons of certain wavelengths can be absorbed by a given kind of atom. ■ Figure 6-4 shows the lowest four energy levels of the hydrogen atom, along with three photons the atom could absorb. The longest-wavelength photon has only enough energy to excite the electron to the second energy level, but the shorter-wavelength photons can excite the electron to higher levels. Because the hydrogen atom has many more energy levels than shown in Figure 6-4, it can absorb photons of many different wavelengths.

Atoms, like humans, cannot exist in an excited state forever. An excited atom is unstable and must eventually (usually within 10^{-9} to 10^{-6} seconds) give up the energy it has absorbed and



■ **Figure 6-4**

A hydrogen atom can absorb only those photons that move the atom's electron to one of the higher-energy orbits. Here, three photons with different wavelengths are shown along with the changes they each would produce in the electron's orbit if they were absorbed.



■ **Figure 6-5**

An atom can absorb a photon only if the photon has the correct amount of energy. The excited atom is unstable and within a fraction of a second returns to a lower energy level, re-radiating the photon in a random direction.

return its electron to a lower energy level. Thus the electron in an excited atom tends to tumble down to its lowest energy level, its ground state.

When an electron drops from a higher to a lower energy level, it moves from a loosely bound level to one that is more tightly bound. The atom then has a surplus of energy—the energy difference between the levels—that it can emit as a photon with a wavelength corresponding to that amount of energy (look back to Chapter 5). Study the sequence of events shown in ■ Figure 6-5 to see how an atom can absorb and emit photons. Because each type of atom or ion has a unique set of energy levels, each type absorbs and emits photons with a unique set of wavelengths. As a result, you can identify the elements in a gas by studying the characteristic wavelengths of light that are absorbed or emitted.

The process of excitation and emission is a common sight in urban areas at night. A neon sign glows when atoms of neon gas in a glass tube are excited by electricity flowing through the tube. As the electrons in the electric current flow through the gas, they collide with the neon atoms and excite them. Almost immediately after a neon atom is excited, its electron drops back to a lower energy level, emitting the surplus energy as a photon of a certain wavelength. The photons emitted by excited neon blend to produce a reddish orange glow. Signs of other colors, generically called "neon signs," contain other gases or mixtures of gases instead of pure neon. Whenever you look at a neon sign, you are seeing atoms absorbing and emitting energy in the form of photons with specific colors determined by the structure of electron orbits in those atoms.

Neon signs are simple, but stars are complex. The colors of stars are not determined by the composition of the gases they contain. In the next section, you will discover why some stars are red and some are blue, and that will give you a new insight into how light interacts with matter.

Radiation from a Heated Object

If you look closely at the stars in the constellation Orion, you will notice that they are not all the same color (look back to Figure 2-4a). One of your Favorite Stars, Betelgeuse, in the upper left corner of Orion, is quite red; another Favorite Star, Rigel, in the lower right corner, is blue. These differences in color arise from differences in temperature.

The starlight that you see comes from gasses that make up the visible surface of the star, its photosphere. (Recall that you first learned about the photosphere of the sun in Chapter 3, in the context of solar eclipses.) Layers of gas deeper inside the star also emit light, but that light is reabsorbed before it can reach the surface. The gas above the photosphere is too thin to emit much light. The photosphere is the visible surface of a star because it is dense enough to emit lots of light but transparent enough to allow that light to escape.

Stars produce their light for the same reason heated horseshoes glow in a blacksmith's forge—because they are hot. If a horseshoe is not too hot, it glows ruddy red, but as it heats up it grows brighter and yellower. Yellow-hot is hotter than red-hot but not as hot as white-hot.

The light from stars and from glowing horseshoes is produced by the acceleration of charged particles. Usually the accelerated particles are electrons because they are the least massive charged particles, and they are on the outsides of atoms, so they are the easiest to get moving. An electron produces a surrounding electric field, and if you disturb an electron, the change in its electric field spreads outward at the speed of light as electromagnetic radiation. You learned in Chapter 4 that "acceleration" means any change in motion. Whenever you change the motion of an electron or other charged particle, you generate electromagnetic waves. If you run a comb through your hair, you disturb electrons in both hair and comb, producing static electricity. That produces electromagnetic radiation, which you can hear as snaps and crackles if you are standing near an AM radio. Stars don't comb their hair, of course, but they are hot, and they are made up of ionized gases, so there are plenty of electrons zipping around.

The molecules and atoms in any object are in constant motion, and in a hot object they are more agitated than in a cool object. You can refer to this agitation as **thermal energy**. If you touch an object that contains lots of thermal energy, it will feel hot as the thermal energy flows into your fingers. The flow of thermal energy is called **heat**. In contrast, **temperature** refers to the average speed of the particles. Hot cheese and hot green beans can have the same temperature, but the cheese can contain more thermal energy and can burn your tongue. Thus, *heat* refers to the flow of thermal energy, and *temperature* refers to the intensity of the agitation among the particles.

When astronomers refer to the temperature of a star, they are talking about the temperature of the gases in the photosphere, and they express those temperatures on the **Kelvin temperature scale**. On this scale, zero degrees Kelvin (written 0 K) is **absolute zero** (−273.2°C or −459.7°F), the temperature at which an object contains no thermal energy that can be extracted. Water freezes at 273 K and boils at 373 K (at sea-level atmospheric pressure). The Kelvin temperature scale is useful in astronomy because it is based on absolute zero and consequently is related directly to the motion of the particles in an object.

Now you can understand why a hot object glows, or to put it another way, why a hot object emits photons, bundles of electromagnetic energy. The hotter an object is, the more motion there is among its particles. The agitated particles, including electrons, collide with each other, and when electrons accelerate—change their motion—part of the energy is carried away as electromagnetic radiation. The radiation emitted by a heated object is called **blackbody radiation**, a name translated from a German term that refers to the way a perfectly opaque object would behave. A perfectly opaque object would be both a perfectly efficient absorber and a perfectly efficient emitter of radiation. At room temperature, such a perfect absorber and emitter would look black, but at higher temperatures it would glow at wavelengths visible to a human eye. That explains why in astronomy and physics contexts you will see the term *blackbody* referring to objects that glow brightly.

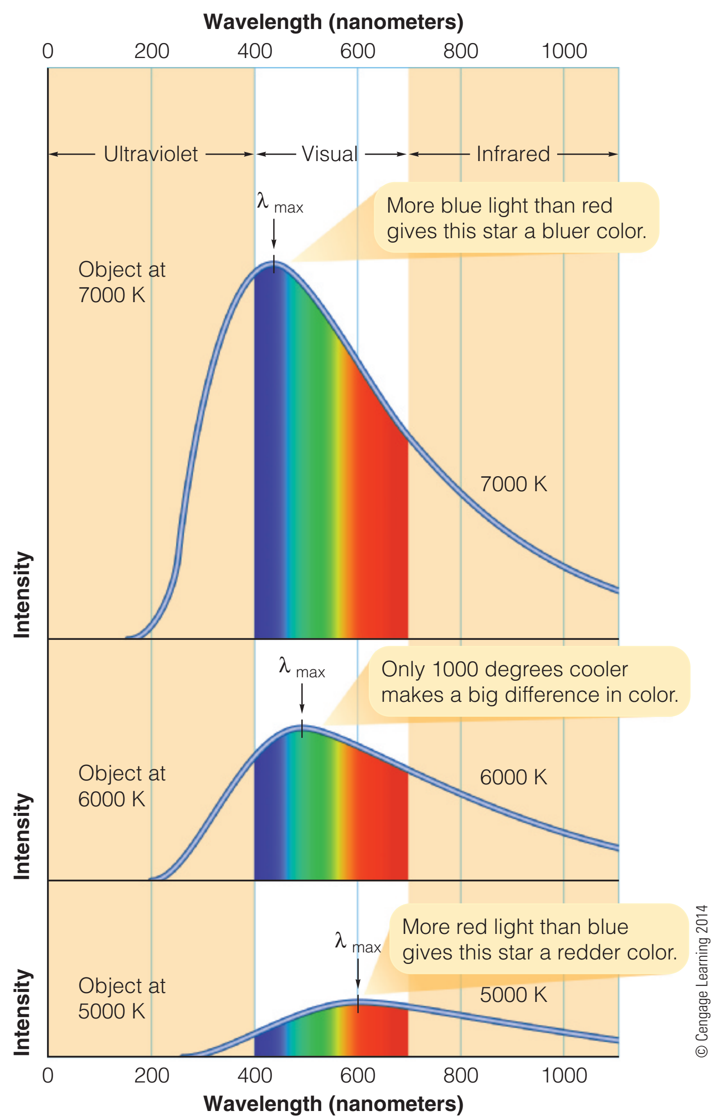
Blackbody radiation is quite common. In fact, it is responsible for the light emitted by an incandescent light bulb. Electricity flowing through the filament of the bulb heats it to high temperature, and it glows. You can also recognize the light emitted by hot lava as blackbody radiation. Many objects in the sky, including the sun and other stars, primarily emit blackbody radiation because they are mostly opaque.

Hot objects emit blackbody radiation, but so do cold objects. Ice cubes are cold, but their temperature is higher than absolute zero, so they contain some thermal energy and must emit some blackbody radiation. The coldest gas drifting in space has a temperature only a few degrees above absolute zero, but it too emits blackbody radiation.

Two features of blackbody radiation are important. First, the hotter an object is, the more blackbody radiation it emits. Hot objects emit more radiation because their agitated particles collide more often and more violently with each other. That's why a glowing coal from a fire emits more total energy than an ice cube of the same size.

The second feature of blackbody radiation is the relationship between the temperature of the object and the wavelengths of the photons it emits. The wavelength of the photon emitted when a particle collides with an electron depends on the violence of the collision. For example, only a violent collision can produce a short-wavelength (high-energy) photon. The electrons in an object have a distribution of speeds; a few move very rapidly, and a few move very slowly, but most travel at intermediate speeds. Because electrons with speeds much higher or much lower than the average speed are rare, extremely violent and extremely gentle collisions don't occur very often, therefore very short-wavelength and very long-wavelength photon emissions are both rare. Consequently, blackbody radiation is made up of photons with a distribution of wavelengths, with medium wavelengths most common. The **wavelength of maximum intensity (λ_{max})** is the wavelength at which the object emits the most intense radiation. Note that λ_{max} does not refer to the maximum wavelength but to the wavelength of the maximum.

■ Figure 6-6 shows the intensity of radiation versus wavelength for three objects of different temperatures. The curves are high in the middle and low at either end because the objects emit



■ **Figure 6-6**

Graphs of blackbody radiation intensity versus wavelength from three objects at different temperatures demonstrate that a hot object radiates more total energy than a cooler object (Stefan-Boltzmann law) and that the wavelength of maximum intensity is shorter for hotter objects than for cooler objects (Wien's law). The hotter object here would look blue to your eyes, while the cooler object would look red.

Reasoning with Numbers | 6-1

Blackbody Radiation

Blackbody radiation is described by two simple laws that can be expressed in precise mathematical form. So many objects give off radiation as blackbodies that these two laws are important principles in the analysis of light from the universe.

Wien's law expresses quantitatively the relation between temperature and the wavelength at which a blackbody radiates the most energy, its wavelength of maximum intensity (λ_{max}). Written for conventional intensity units, the law is:

$$\lambda_{max}(\text{nm}) = \frac{2.90 \times 10^6}{T\,(\text{K})}$$

That is, the wavelength in nanometers of maximum radiation intensity emitted by a blackbody equals 2.9 million divided by the blackbody's temperature on the Kelvin scale.

This law is a powerful aid in astronomy. For example, a cool star with a temperature of 2900 K will emit most intensely at a wavelength of 1000 nm, which is in the near-infrared part of the spectrum. In contrast, a very hot star, with a temperature of 29,000 K, radiates most intensely at a wavelength of 100 nm, which is in the ultraviolet.

The Stefan-Boltzmann law relates the temperature of a blackbody to the total radiated energy. Recall from Chapter 5 that energy is expressed in units called joules (symbolized by capital J). (The energy of motion of an apple hitting the floor after falling from a table is about one joule.) The total radiation in units of joules per second given off by 1 square meter of the surface of the object equals a constant number, represented by the Greek lower case sigma (σ), multiplied by the temperature raised to the fourth power:*

$$E = \sigma T^4 \ (\text{J/s/m}^2)$$

How does this help you understand stars? Suppose a star the same size as the sun has a surface temperature twice as hot as the sun's surface. Then each square meter of that star radiates not just twice as much energy, but 2^4, or 16, times as much energy. From this law you can see that a small difference in temperature between two stars can produce a large difference in the amount of energy emitted from the stars' surfaces.

*For the sake of completeness, you can note that the constant σ equals 5.67×10^{-8} J/(s m^2 K^4) (units of Joules per second per square meter per degree Kelvin to the fourth power).

radiation most intensely at intermediate wavelengths. The total area under each curve is proportional to the total energy emitted, and you can see that the hotter object emits more total energy than the cooler objects. This rule is known as the **Stefan-Boltzmann law**, named after the two physicists who discovered it.

Look again closely at the curves in Figure 6-6, and you will see that the wavelength of maximum intensity depends on temperature. A hotter object will have a shorter wavelength of maximum emitted intensity. The figure thus shows how temperature determines the color of a glowing blackbody. The hotter object emits more blue light than red and thus looks blue, and the cooler object emits more red than blue and consequently looks red. This rule is known as **Wien's law**. Now you can understand why the two Favorite Stars in Orion mentioned previously, Betelgeuse and Rigel, have such different colors. Betelgeuse is relatively cool and so looks red, but Rigel is hot and looks blue.

You can see both Wien's law and the Stefan-Boltzmann law in operation if you look down into a toaster after you start the toast. You first see a faint deep red glow that becomes brighter (Stefan-Boltzmann law) and more orange-yellow in color (Wien's law) as the coils get hotter. **Reasoning with Numbers 6-1** gives you some practice at understanding the behavior of blackbody emission.

Objects much cooler than stars don't glow at visible wavelengths but still produce blackbody radiation. For example, the human body has a temperature of 310 K and emits blackbody radiation mostly in the infrared part of the spectrum. Infrared security cameras can detect burglars by the radiation they emit, and mosquitoes can track you down in total darkness by homing in on your infrared radiation. Although you emit lots of infrared radiation, you rarely emit higher-energy photons, and you almost never emit an X-ray or gamma-ray photon. The wavelength of maximum intensity of your glow lies in the infrared part of the spectrum.

SCIENTIFIC ARGUMENT

The infrared radiation coming out of your ear can tell a doctor your temperature. How does that work?

You know two radiation laws, so your argument must use the right one. Doctors and nurses use a handheld device to measure body temperature by observing the infrared radiation emerging from a patient's ear. You might suspect the device depends on the Stefan-Boltzmann law and therefore measures the intensity of the infrared radiation. It is true that a person with a fever will emit more energy than a healthy person. However, a healthy person with a large ear canal would emit more than a person with a small ear canal, so measuring intensity would not necessarily be helpful. The medical device actually depends on Wien's law, finding temperature by measuring the "color" of the infrared radiation. A patient with a fever will emit at a slightly shorter wavelength of maximum intensity, and the infrared radiation emerging from his or her ear will be a tiny bit "bluer" than that emitted by a person with normal temperature.

Astronomers can measure the temperatures of stars the same way. Adapt your argument for stars. **Use Figure 6-6 to explain how the colors of stars reveal their temperatures.**

6-3 Understanding Spectra

SCIENCE IS A WAY OF UNDERSTANDING NATURE, and the spectrum of a star can tell you a great deal about the star's temperature, motion, and composition. In later chapters, you will use spectra to study other astronomical objects such as galaxies and planets, but you can begin by looking at the spectra of stars, including that of the sun.

The spectrum of a star is formed as light passes outward through the gases near its surface. Read **Atomic Spectra** on pages 110–111 and notice that it describes three important properties of spectra and defines 12 new terms that will help you understand astronomical spectra:

1. There are three kinds of spectra: (i) *continuous spectra;* (ii) *absorption* or *dark-line spectra,* which contain *absorption lines;* and (iii) *emission* or *bright-line spectra,* which contain *emission lines.* These kinds of spectra are described by *Kirchhoff's laws.* When you see one of these kinds of spectra, you can recognize the arrangement of matter that emitted the light.
2. Photons are emitted or absorbed when an electron in an atom makes a *transition* from one energy level to another. The wavelengths of the photons depend on the energy difference between the two levels, so each spectral line represents not one energy level but rather an electron jump between two energy levels. Hydrogen atoms can produce many spectral lines that are grouped in series such as the *Lyman, Balmer,* and *Paschen* series. Only three hydrogen lines, all in the Balmer series, are visible to human eyes. The emitted photons coming from a hot cloud of hydrogen gas have the same wavelengths as the photons absorbed by hydrogen atoms in the gases of a star.
3. Most modern astronomy books display spectra as graphs of intensity versus wavelength. Be sure you recognize the connection between the dark absorption lines and the dips in the graphed spectrum.

Imagine you are an astronaut with a hand-held spectrograph, approaching a fresh lava flow on a moon with no atmosphere. You aim your spectrograph straight at the lava flow; what kind of spectrum do Kirchhoff's laws say you will see? You'll get a continuous (blackbody) spectrum, with all the colors of the rainbow present, given off by the opaque, glowing hot lava. And, as a bonus, measuring the wavelength of the strongest blackbody emission lets you determine the temperature of the lava using Wien's law.

Suddenly, the lave flow begins to bubble, and gas trapped in the molten rock is released, making a temporary warm, thin atmosphere right above the lava flow. If you point your spectrograph to look through that gas with the hot lava as the background, you will observe an absorption (dark line) spectrum in which the lava's blackbody spectrum is now interrupted by missing colors caused by atoms in the gas absorbing photons on their way from the lava to you. Finally, you crouch down and point your spectrograph at the gas at such an angle that the background is not hot lava but cold, empty dark sky. Now, you will see an emission (bright-line) spectrum, produced by atoms in the gas releasing photons as their electrons drop down toward the ground state, with lines at the same wavelengths as in the absorption spectrum of the same gas.

Chemical Compositions

Identifying the elements that are present in a star, planet, or gas cloud by identifying the lines in that object's spectrum is a relatively straightforward procedure. For example, two dark absorption lines appear in the yellow region of the sun's spectrum at the wavelengths 589.0 nm and 589.6 nm. The only atom that can produce this pair of lines is sodium, so the sun must contain sodium. Over 90 elements in the sun have been identified this way (■ Figure 6-7a).

However, just because the spectral lines that identify an element are missing, you cannot conclude that the element itself is absent. For example, the spectral lines in the hydrogen Balmer series are weak in the sun's spectrum, even though 90 percent of the atoms in the sun are hydrogen. A later chapter will discuss the discovery that this occurs because the sun is too cool to produce strong hydrogen Balmer lines. Astronomers must consider that an element's spectral lines may be absent from an object's spectrum not because that element is missing but because that object has the wrong temperature to excite those atoms to the energy levels that produce visible spectral lines.

To derive accurate chemical abundances, astronomers must use the physics that describes the interaction of light and matter to analyze a spectrum, take into account the object's temperature, and calculate the amounts of the elements present there. Such results show that nearly all stars, and most of the visible matter in the universe, have a chemical composition similar to the sun's—about 91 percent of the atoms are hydrogen, and 8.9 percent are helium, with small traces of heavier elements. You will use these results in later chapters when you study the life stories of the stars, the history of our galaxy, and the origin of the universe.

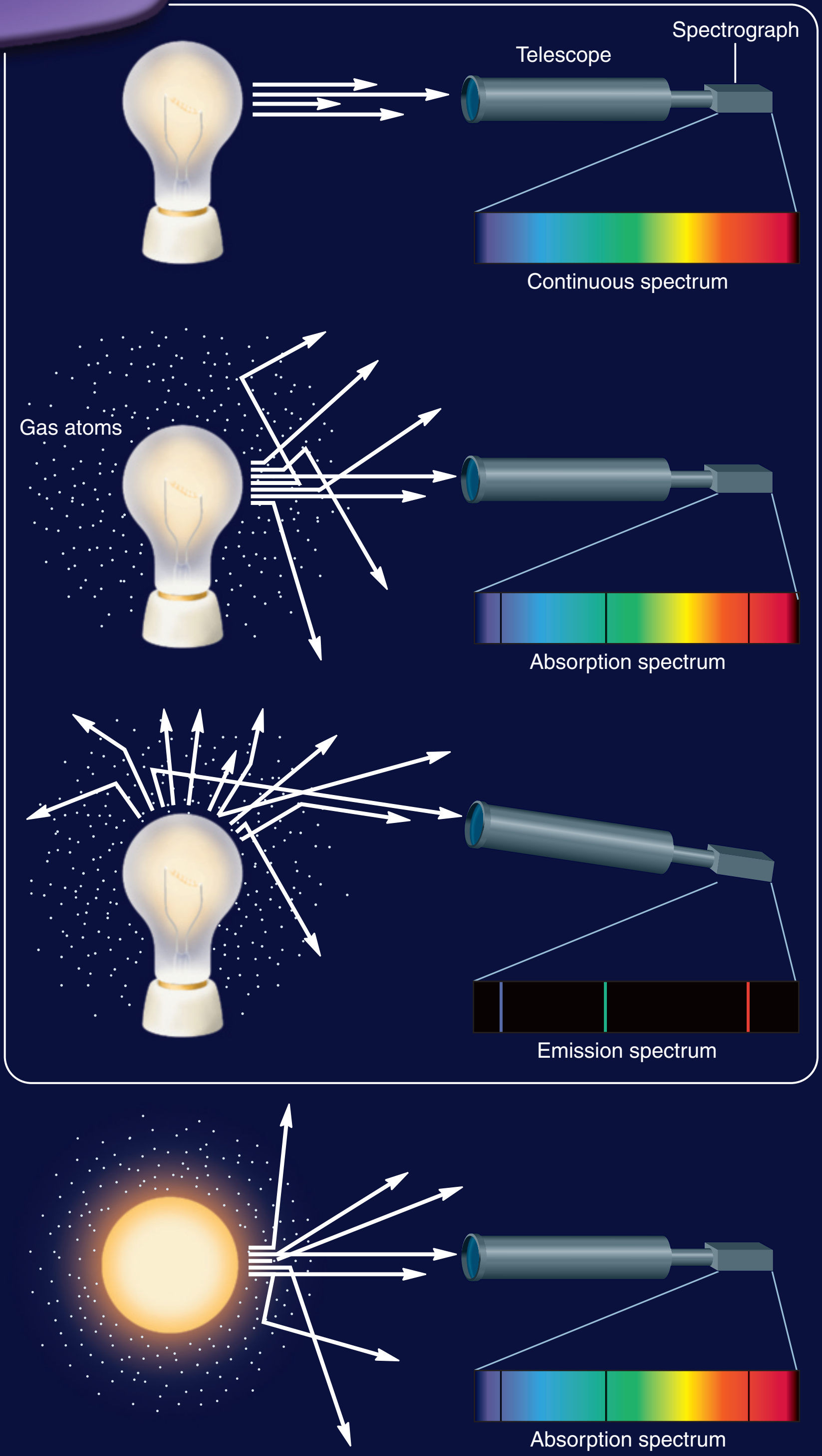
The Doppler Effect

Surprisingly, one of the pieces of information hidden in a spectrum is the velocity of the light source. Astronomers can measure the wavelengths of the lines in a star's spectrum and find the velocity of the star. The **Doppler effect** is the apparent change in the wavelength of radiation caused by the motion of the source.

When astronomers talk about the Doppler effect, they are talking about a shift in the wavelength of electromagnetic radiation. But the Doppler shift can occur in any form of wave phenomena, including sound waves. You probably hear the Doppler effect every day without noticing.

The pitch of a sound is determined by its wavelength. Sounds with long wavelengths have low pitches, and sounds with

Atomic Spectra



1 To understand how to analyze a spectrum, begin with a simple incandescent lightbulb. The hot filament emits blackbody radiation, which forms a **continuous spectrum.**

An **absorption spectrum** results when radiation passes through a cool gas. In this case you can imagine that the lightbulb is surrounded by a cool cloud of gas. Atoms in the gas absorb photons of certain wavelengths, which are missing from the spectrum, and you see their positions as dark **absorption lines.** Such spectra are sometimes called **dark-line spectra.**

An **emission spectrum** is produced by photons emitted by an excited gas. You could see **emission lines** by turning your telescope aside so that photons from the bright bulb did not enter the telescope. The photons you would see would be those emitted by the excited atoms near the bulb. Such spectra are also called **bright-line spectra.**

1a The spectrum of a star is an absorption spectrum. The denser layers of the photosphere emit blackbody radiation. Gases in the atmosphere of the star absorb their specific wavelengths and form dark absorption lines in the spectrum.

Fair Use.

1b In 1859, long before scientists understood atoms and energy levels, the German scientist Gustav Kirchhoff formulated three rules, now known as **Kirchhoff's laws,** that describe the three types of spectra.

KIRCHHOFF'S LAWS

Law I: The Continuous Spectrum

A solid, liquid, or dense gas excited to emit light will radiate at all wavelengths and thus produce a continuous spectrum.

Law II: The Emission Spectrum

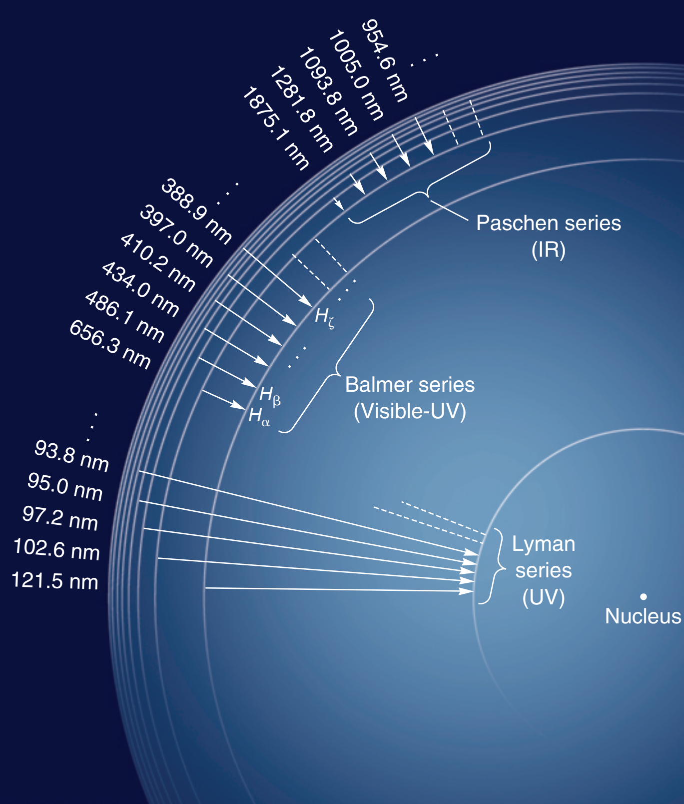
A low-density gas excited to emit light will do so at specific wavelengths and thus produce an emission spectrum.

Law III: The Absorption Spectrum

If light comprising a continuous spectrum passes through a cool, low-density gas, the result will be an absorption spectrum.

NOAO/AURA/NSF

Line art on this page © Cengage Learning 2014



2 The electron orbits in the hydrogen atom are shown here as energy levels. When an electron makes a **transition** from one orbit to another, it changes the energy stored in the atom. In this diagram, arrows pointed inward represent transitions that result in the emission of a photon. If the arrows pointed outward, they would represent transitions that result from the absorption of a photon. Long arrows represent large amounts of energy and correspondingly short-wavelength photons.

2a Transitions in the hydrogen atom can be grouped into series—the **Lyman series, Balmer series, Paschen series,** and the like. Transitions and the resulting spectral lines are identified by Greek letters. Only the first few transitions in the first three series are shown at left.

2b In this drawing (right) of the hydrogen spectrum, emission lines in the infrared and ultraviolet are shown as gray. Only the first three lines of the Balmer series are visible to human eyes.

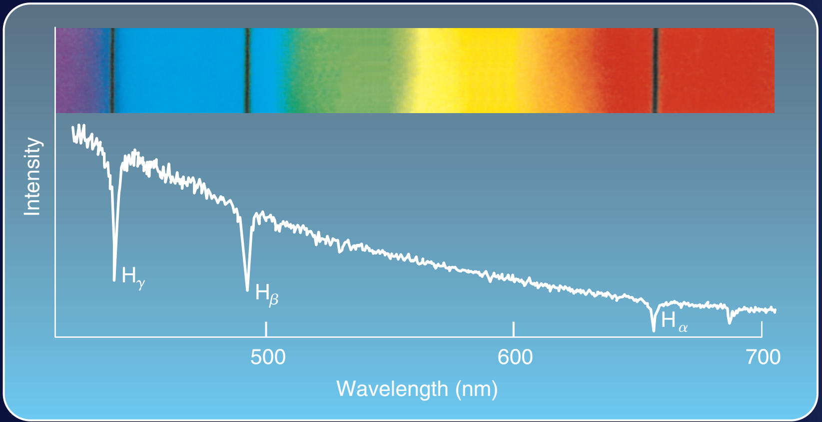

Visual-wavelength image

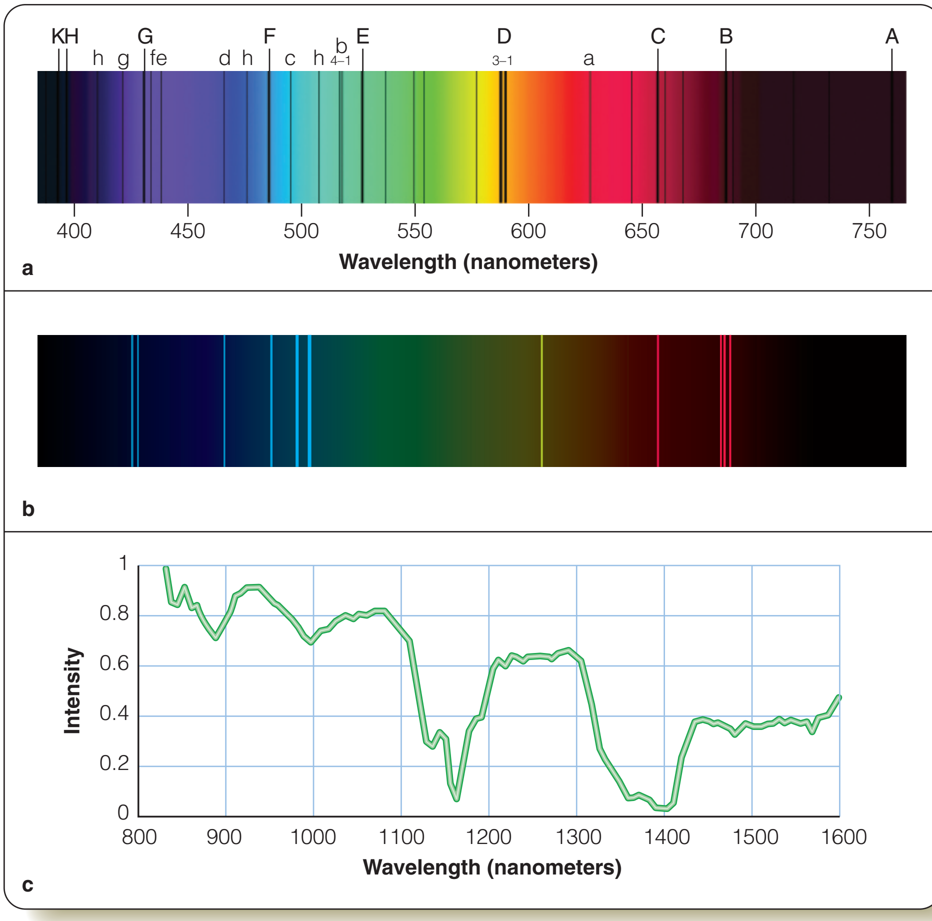
AURA/NOAO/NSF

2c Excited clouds of gas in space emit light at all of the Balmer wavelengths, but you see only the red, blue, and violet photons blending to create the pink color typical of ionized hydrogen.

The shorter-wavelength lines in each series blend together.

3 Modern astronomers rarely work with spectra as bands of light. Spectra are usually recorded digitally, so it is easy to represent them as graphs of intensity versus wavelength. Here, the artwork above the graph suggests the appearance of a stellar spectrum. The graph below reveals details not otherwise visible and allows comparison of relative intensities. Notice that dark absorption lines in the spectrum appear as dips in the curve of intensity.





Courtesy Maureen V. Gebruiker

John Talbot/LaserStars.org

© ESA/NASA/LPL, U. Arizona/Observatoire de Paris-Meudon

Figure 6-7

(a) The sun's spectrum at visual wavelengths. The bright colored background shows the continuous spectrum of blackbody emission from the sun's photosphere. The dark spectral absorption lines represent precise colors (photons of exact energies) removed from the sun's radiation by atoms in its transparent atmosphere. (b) A model of the visual wavelength emission (bright-line) spectrum of NGC 2392, the nebula pictured in Figure 6-1. Emission lines from ionized atoms of hydrogen (red), nitrogen (red) and oxygen (green), among others, are seen. (c) Near-infrared spectrum of the atmosphere and surface of Saturn's moon Titan measured by the Huygens probe at an altitude of 20 meters (65 ft) using a light source on the bottom of the probe.

short wavelengths have higher pitches. Every time a car or truck passes you and the pitch of its engine noise seems to drop, you are hearing the Doppler effect. The vehicle's sound is shifted to shorter wavelengths and higher pitches while it is approaching and is shifted to longer wavelengths and lower pitches after it passes.

To see why the sound waves are shifted in wavelength, consider a fire truck approaching you with a bell clanging once a second. When the bell clangs, the sound travels ahead of the truck to reach your ears. One second later, the bell clangs again, but, during that one second, the fire truck has moved closer to you, so the bell is closer at its second clang. Now the sound has a shorter distance to travel and reaches your ears a little sooner than it would have if the fire truck were not approaching. If you timed the clangs, you would find that you heard them slightly less than one second apart. After the fire truck passes you and is moving away, you hear the clangs sounding slightly more than one second apart because now each successive clang of the bell occurs farther from you and the sound travels farther to reach your ears.

Figure 6-8a shows a fire truck moving toward one observer and away from another observer. The position of the bell at each clang is shown by a small black bell. The sound of the clangs spreading outward is represented by black circles. You can see how the clangs are squeezed together ahead of the fire truck and stretched apart behind. If the fire truck has a siren instead of a bell, the sound coming from the siren will be a wave that can be depicted as a series of peaks and valleys representing compressions and uncompressions of the air. If the truck and siren are moving toward an observer, the peaks and valleys of the siren's sound wave will arrive more often—at a higher frequency—than if the truck were not moving, and the observer will hear the siren at a higher pitch than the same siren when it is stationary.

Now you can substitute a source of light for the clanging bell or wailing siren (Figure 6-8b). Imagine the light source emitting waves continuously as it approaches you. Each time the source emits the peak of a wave, it will be slightly closer to you than when it emitted the peak of the previous wave. From your vantage point, the successive peaks of the wave will seem closer together in the same way that the clangs of the bell seemed closer together. The light will appear to have a shorter wavelength, making it slightly bluer. Because the light is shifted slightly toward the blue end of the spectrum, this is called a **blueshift**. After the light source has passed you and is moving away, the peaks of successive waves seem farther apart, so the light has a longer wavelength and is redder. This is a **redshift**.

The terms *redshift* and *blueshift* are used to refer to any range of wavelengths. The light does not actually have to be red or blue, and the terms apply equally to wavelengths in other parts of the electromagnetic spectrum such as X-rays and radio waves. *Red* and *blue* refer to the direction of the shift, not to actual color.

The amount of change in wavelength, and thus the size of the Doppler shift, depends on the velocity of the source. A moving car has a smaller Doppler shift than a jet plane, and a slow-moving star has a smaller Doppler shift than one that is moving more quickly. You can measure the velocity of a star by measuring the size of its Doppler shift. If a star is moving toward Earth, it has a blue shift and each of its spectral lines is shifted very slightly toward shorter wavelengths. If it is receding from Earth, it has a red shift. The shifts are much too small to change the color of a star, but they are easily detected in spectra. **Reasoning with Numbers 6-2** shows you how astronomers can convert Doppler shifts into velocities.

When you think about the Doppler effect, it is important to understand two things. Earth itself moves, so a measurement of a Doppler shift really measures the relative motion between Earth and the star. Figure 6-8c shows the Doppler effect in two spectra of the star Arcturus. Lines in the top spectrum are slightly blueshifted because the spectrum was recorded when Earth, in the course of its orbit, was moving toward Arcturus. Lines in the bottom spectrum are redshifted because it was recorded six months later, when Earth was moving away from Arcturus.

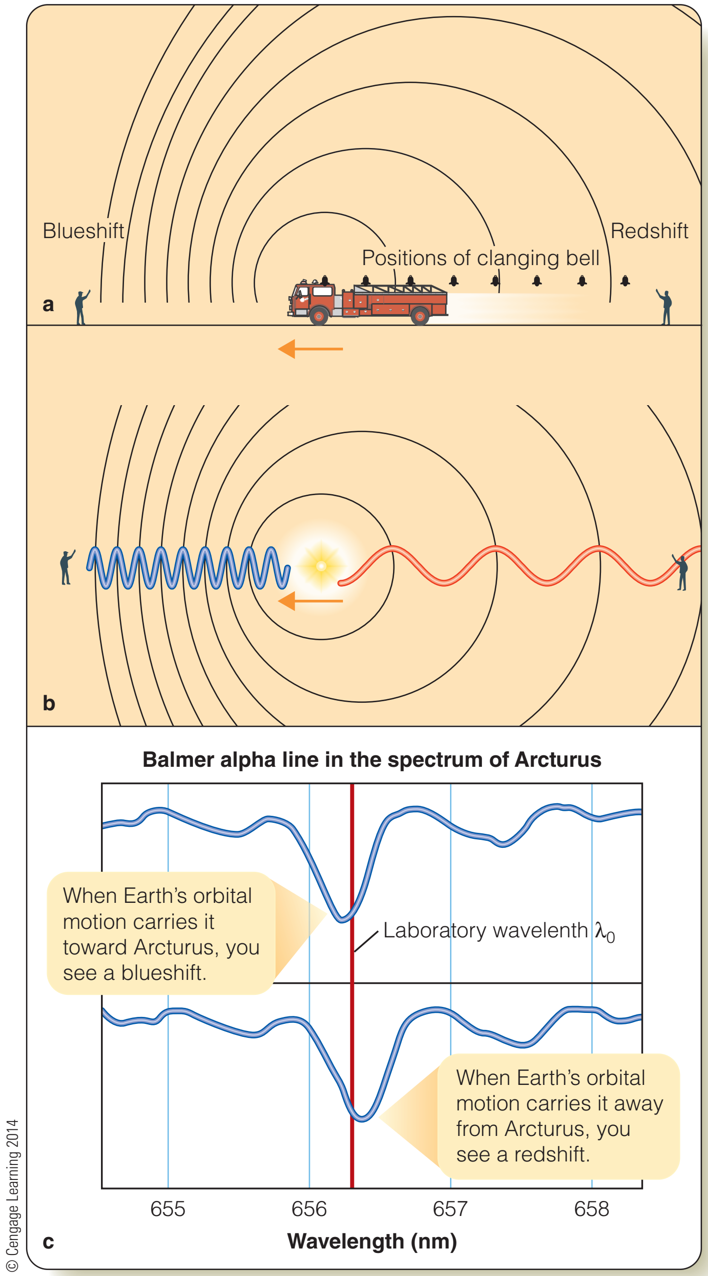


Figure 6-8

The Doppler effect. (a) The clanging bell on a moving fire truck produces sounds that move outward (black circles). An observer ahead of the truck hears the clangs closer together, while an observer behind the truck hears them farther apart. Similarly, the sound waves from a siren on an approaching truck will be received more often, and thus be heard with a higher pitch, than a stationary truck, and the siren will have a lower pitch if it is going away. (b) A moving source of light emits waves that move outward (black circles). An observer toward whom the light source is moving observes a shorter wavelength (a blueshift), and an observer for whom the light source is moving away observes a longer wavelength (a redshift). (c) Absorption lines in the spectrum of the bright star Arcturus are shifted toward the blue in winter, when Earth's orbital motion carries it toward the star, and toward the red in summer when Earth moves away from the star.

Reasoning with Numbers | 6-2

The Doppler Formula

Astronomers can measure radial velocity by using the Doppler effect. The "laboratory" wavelength λ_0 is the wavelength a certain spectral line would have if the source of the light is not moving relative to the spectrograph. In the spectrum of a star, this spectral line is shifted by some small amount, $\Delta\lambda$. If the wavelength is increased (a red shift), $\Delta\lambda$ is positive; if the wavelength is decreased (a blue shift), $\Delta\lambda$ is negative. The radial velocity, V_r, of the star is given by the Doppler formula:

$$\frac{V_r}{c} = \frac{\Delta\lambda}{\lambda_0}$$

That is, the radial velocity divided by the speed of light, c, is equal to $\Delta\lambda$ divided by λ_0. In astronomy, radial velocities are almost always given in kilometers per second, so c is expressed as 300,000 km/s.

For example, suppose the laboratory wavelength of a certain spectral line is 600.00 nm, and the line is observed in a star's spectrum at a wavelength of 600.10 nm. Then $\Delta\lambda$ is $+0.10$ nm, and the radial velocity is 0.10/600 multiplied by the speed of light, which equals 50 km/s. Because $\Delta\lambda$ is positive, you know the star is receding from you.

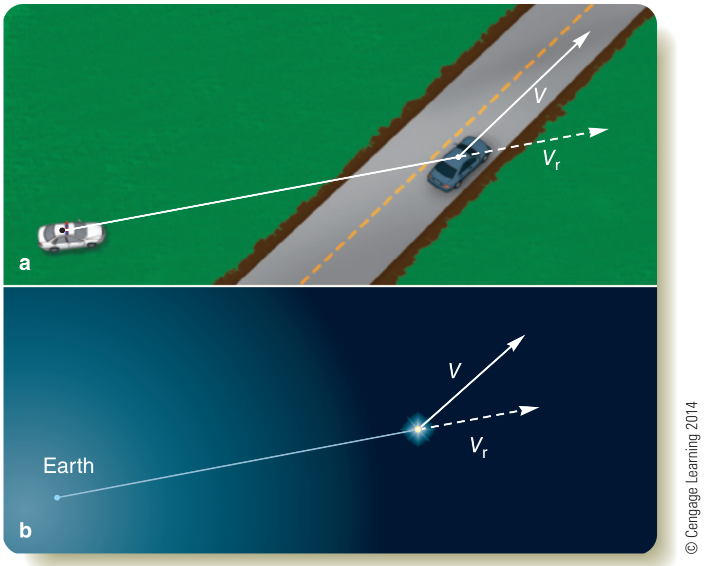


Figure 6-9

(a) Police radar can measure only the radial part of your velocity *(V_r)* as you drive down the highway, not your true velocity along the pavement *(V)*. That is why police using radar should never park far from the highway. This police car is poorly placed to make a good measurement. (b) From Earth, astronomers can use the Doppler effect to measure the radial velocity *(V_r)* of a star, but they cannot measure its true velocity, *V*, through space.

To find the true motion of Arcturus, astronomers must account for the motion of Earth.

The second point to remember is that the Doppler shift is sensitive only to the part of the velocity directed away from you or toward you—the **radial velocity** *(V_r)* (Figure 6-9). You cannot use the Doppler effect to detect any part of the velocity that is perpendicular to your line of sight. A star moving to the left (Figure 6-9b) would have no blueshift or redshift because its distance from Earth would not be decreasing or increasing. This is why police using radar guns should park right next to the highway (Figure 6-9a). They want to measure your full velocity as you drive toward them, not just part of your velocity.

Armed with your new understanding of light and spectra, you are ready to focus on your first astrophysical object, the star that supports life on Earth, the sun.

What Are We? Stargazers

Do you suppose chickens ever look at the sky and wonder what the stars are? Probably not. Chickens are very good at the chicken business, but they are not known for big brains and deep thought. Humans, in contrast, have highly evolved, sophisticated brains and are extremely curious. In fact, curiosity may be the most reliable characteristic of intelligence, and curiosity about the stars is a natural extension of our continual attempts to understand the world around us.

For early astronomers like Copernicus and Kepler, the stars were just points of light. There seemed to be no way to learn anything about them. Galileo's telescope revealed surprising details about the planets; but, even viewed through a large telescope, the stars are just points of light. When later astronomers began to realize that stars were other suns, the stars still seemed forever beyond human knowledge.

As you have seen, the key is to understanding the universe is knowing how light interacts with matter. In the last 150 years or so, scientists have discovered how atoms and light interact to produce the spectra we observe, and astronomers have applied those discoveries to the ultimate object of human curiosity—the stars.

Chickens may never wonder what the stars are, or even wonder what chickens are, but humans are curious animals, and we do wonder about the stars and about ourselves. Our yearning to understand the stars is just part of our quest to understand what we are.

Study and Review

Summary

- Modern astronomy is more properly called **astrophysics (p. 101)**, a field of study that interprets astronomical observations in terms of physics theory and laboratory experiments in order to understand the compositions, internal processes, and histories of celestial objects.
- An atom consists of a **nucleus (p. 102)** surrounded by a cloud of **electrons (p. 102).** The nucleus is made up of positively charged **protons (p. 102)** and uncharged **neutrons (p. 102).**
- The number of protons in an atom determines which element it is.
- Atoms of the same element (that is, having the same number of protons) with different numbers of neutrons are called **isotopes (p. 103).**
- A neutral atom is surrounded by a number of negatively charged electrons equal to the number of protons in the nucleus. An atom that has lost or gained an electron is said to be **ionized (p. 103)** and is called an **ion (p. 103).**
- Two or more atoms joined together form a **molecule (p. 103).**
- The electrons in an atom are attracted to the nucleus by the **Coulomb force (p. 103).** As described by **quantum mechanics (p. 104)**, the **binding energy (p. 103)** that holds electrons in an atom is limited to certain energies, and that means the electrons may occupy only certain **permitted orbits (p. 104).**
- The size of an electron's orbit depends on its energy, so the orbits can be thought of as **energy levels (p. 105)**, with the lowest possible energy level known as the **ground state (p. 105).**
- An **excited atom (p. 105)** is one in which an electron is raised to a higher orbit by a collision between atoms or the absorption of a photon of the proper energy.
- The agitation among the atoms and molecules of an object is called **thermal energy (p. 106)**, and the flow of thermal energy is **heat (p. 106).** In contrast, **temperature (p. 106)** refers to the intensity of the agitation and can be expressed on the **Kelvin temperature scale (p. 107)**, which gives temperature above **absolute zero (p. 107).**
- Collisions among the particles in a body accelerate electrons and cause the emission of **blackbody radiation (p. 107).** The hotter an object is, the more total energy it radiates (known as the **Stefan-Boltzmann law (p. 108)**, and the shorter is its **wavelength of maximum intensity, λ_{max} (p. 107)** (known as **Wien's law, p. 108**). This allows astronomers to estimate the temperatures of stars from their colors.
- **Kirchhoff's laws (p. 110)** summarize how (1) a hot solid, liquid, or dense gas emits electromagnetic radiation at all wavelengths and produces a **continuous spectrum (p. 110);** (2) an excited low-density gas produces an **emission (bright-line) spectrum (p. 110)** containing **emission lines (p. 110);** and (3) a light source viewed through a low-density gas produces an **absorption (dark-line) spectrum (p. 110)** containing **absorption lines (p. 110).**
- An atom can emit or absorb a photon when an electron makes a **transition (p. 111)** between orbits.
- Because orbits of only certain energy differences are permitted in an atom, photons of only certain wavelengths can be absorbed or emitted. Each kind of atom has its own characteristic set of spectral lines. The hydrogen atom has the **Lyman series (p. 111)** of lines in the ultraviolet, the **Balmer series (p. 111)** partially in the visible, and the **Paschen series (p. 111)** (plus others) in the infrared.
- A spectrum can tell you the chemical composition of the stars. The presence of spectral lines of a certain element shows that that element must be present in the star, but you must proceed with care. Lines of a certain element may be weak or absent if the star is too hot or too cool even if that element is present in the star's atmosphere.
- The **Doppler effect (p. 109)** can provide clues to the motions of the stars. When a star is approaching, you observe slightly shorter wavelengths, a **blueshift (p. 112)**, and when it is receding, you observe slightly longer wavelengths, a **redshift (p. 112).** This Doppler effect reveals a star's **radial velocity, V_r (p. 114)**, that part of its velocity directed toward or away from Earth.

Review Questions

1. Why might you say that atoms are mostly empty space?
2. What is the difference between an isotope and an ion?
3. Why is the binding energy of an electron related to the size of its orbit?
4. Explain why ionized calcium can form absorption lines, but ionized hydrogen cannot.
5. Describe two ways an atom can become excited.
6. Why do different atoms have different lines in their spectra?
7. Why does the amount of blackbody radiation emitted depend on the temperature of the object?
8. Why do hot stars look bluer than cool stars?
9. What kind of spectrum does a neon sign produce?
10. Why does the Doppler effect detect only radial velocity?
11. How can the Doppler effect explain shifts in both light and sound?
12. **How Do We Know?** How is the world you see around you determined by a world you cannot see?

Discussion Questions

1. In what ways is the model of an atom a scientific model? In what ways is it incorrect?
2. Before Fraunhofer and others worked to observe and interpret spectra, most people were of the opinion that we would never know the composition of the sun and stars. Can you think of any scientific question today that most people believe will probably never be answered?

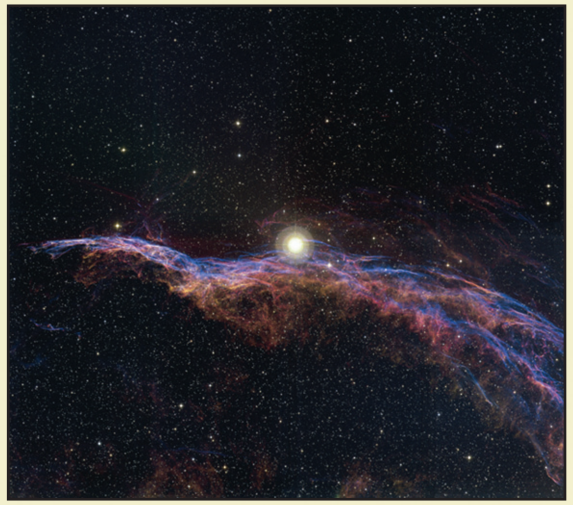
Problems

1. Human body temperature is about 310 K (3.10×10^2 K, or 98.6°F). At what wavelength do humans radiate the most energy? What kind of radiation do we emit?
2. If a star has a surface temperature of 20,000 K (2.00×10^4 K), at what wavelength will it radiate the most energy?
3. Infrared observations of a star show that it is most intense at a wavelength of 2000 nm (2.00×10^3 nm). What is the temperature of the star's surface?
4. If you double the temperature of a blackbody, by what factor will the total energy radiated per second per square meter increase?
5. If one star has a temperature of 6000 K and another star has a temperature of 7000 K, how much more energy per second will the hotter star radiate from each square meter of its surface?

6. Electron orbit transition *A* produces light with a wavelength of 500 nm. Transition *B* involves twice as much energy as *A*. What wavelength is the light it produces?
7. In a laboratory, the Balmer beta line has a wavelength of 486.1 nm. If the line appears in a star's spectrum at 486.3 nm, what is the star's radial velocity? Is it approaching or receding?
8. The highest-velocity stars an astronomer might observe in the Milky Way Galaxy have velocities of about 400 km/s (4.00×10^2 km/s). What change in wavelength would this cause in the Balmer-beta line (laboratory wavelength given in Problem 7)?

Learning to Look

1. Consider Figure 6-3. When an electron in a hydrogen atom moves from the third orbit to the second orbit, the atom emits a Balmer-alpha photon in the red part of the spectrum. In what part of the spectrum would you look to find the photon emitted when an electron in a helium atom makes the same transition?
2. Where should the police car in Figure 6-9 have parked to make a good measurement?
3. The nebula shown below contains mostly hydrogen excited to emit photons. What kind of spectrum would you expect this nebula to produce?



T. Rector, University of Alaska, and WIYN/NURO/AURA/NSF

4. If the nebula crosses in front of the star, and the nebula and star have different radial velocities, what might the spectrum of the star look like?

Great Debates

1. *What will you die into?* Your body is made of matter and energy. Scientists understand that when you die your matter and energy are returned to the universe. Some companies provide policies that let you choose your next life form and guarantee your choice. Are the policies issued by these companies moral and ethical? If you could choose, what object would you like your matter to be next? How would you counsel a friend who wants to purchase their next life form from one of these companies?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

2. *Are Ghosts Blackbodies?* A living human generates energy, but you cannot see this energy. You can test this hypothesis by turning out all lights. Can you see another human in the room? If the room has no light, you cannot see another human. Energy cannot be created or destroyed, but it can be converted from one type to another. With this in mind, when a human dies, where does the human's thermal and chemical energy go? Suppose your energy can be maintained in the form of a ghost. Is a ghost a blackbody, and, if so, in which band of the electromagnetic spectrum would the ghost mostly emit?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

3. *Red and Blue: Part 1.* Because oceans are blue and fire is red, humans may associate the color blue with cool and red with hot. You may have come into this class with this internal bias. As you learned in this chapter, red light is emitted by cooler blackbodies, and blue light is emitted by hotter blackbodies, or the opposite of human perception. You may have noticed that NASA falsely colors some images with blue indicating a cool color and red indicating a hot color. For example, the Wilkinson Microwave Anisotropy Probe (WMAP) shows the Milky Way Galaxy running horizontally across the middle of the page, colored red. In the image, the red color represents regions in space that are 0.0002 K hotter than the blue color. Is falsely coloring images red to represent hot and blue to represent cool doing a disservice to the general public? Should the colors in images like WMAP be changed? Do you think the general public should be educated to the true physics in these public, press-release images?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

4. *Red and Blue: Part 2.* Read about falsely colored pubic images like WMAP in the previous question. Now that your internal bias has been corrected on this subject, do you think you should correct a child who may choose a red color for the sun and a blue color for water and sky when drawing a sunrise over the ocean?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

5. *The Eskimo Nebula.* The Eskimo Nebula, shown in Figure 6-1, got its name because it resembles a fur parka around a face when viewed through a telescope. The use of the word *Eskimo* is now considered politically incorrect in Canada and Greenland, and the indigenous names of specific tribes are preferred. In astronomy, names do not usually change, even if they are incorrect. Should the name be stricken and the object referred to as NGC 2392 or by its other name, the Clownface Nebula? Is the name politically incorrect?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Bohr Atom
- The Doppler Effect

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 2: Light and Matter

If you want to understand how stars evolve, you need to know how they emit energy. This chapter offers you the insight that stars and planets are opaque glowing objects termed "blackbodies" that follow the same rules as do opaque glowing objects in laboratories on Earth. How much energy a blackbody such as a star emits, and at what wavelengths, depends on its temperature.

In 1835, natural philosopher Auguste Comte said that one thing is for certain: We'll never know the composition of the stars. He was proved wrong in less than 50 years. From experiments in laboratories, physicists learned about the structure of atoms and how each type of atom interacts in characteristic ways with electromagnetic radiation. Because atoms in distant stars are the same kinds of atoms as atoms on Earth and follow the same natural laws, we can in fact know the composition of the stars.

Sections 2, 3, and 4 of Virtual Astronomy Lab 2, "Light and Matter," continue your study of interactions of matter with radiation begun in this chapter and the previous chapter. For example, this lab contains demonstrations of Wien's law of radiation. You should note that natural laws are the most fundamental kinds of scientific knowledge. They have been tested over and over so many times scientists have great confidence that they are a true description of how nature works, not only on Earth where they were discovered but throughout the universe as well. You've heard of Newton's law of gravity; now you can add Wien's law of radiation to your list of natural laws that describe the universe. You can access that lab at http://cengage.com/someaddress/VAL/Lab2.

Virtual Astronomy Lab 3: The Doppler Effect

You experience the Doppler effect every day, probably many times a day. Each time a car passes, you hear a whoosh that is higher pitched at the start and lower pitched at the end of the encounter. That is the Doppler effect manifested in sound—you receive the sound waves from the car as it approaches you with higher frequency (pitch) and shorter wavelength than when they left that source, and with lower frequency (pitch) and longer wavelength as it recedes. At the moment of closest approach, when its motion is neither toward nor away from you, the sound waves are received with the same frequency and wavelength as emitted.

Decades ago there was a clever and entertaining demonstration of the Doppler effect arranged by the late Professor Philip Morrison of MIT for a TV science show. A brass band, dressed in tuxedos, stood on a railroad flat car playing one note as the train pulled them past the film crew standing in a station. The band's note was sharp as they approached the observer (film crew), on pitch as the car passed, and flat as they continued away down the tracks. This was contrasted with a recording made by someone standing with the band on the train, verifying that the band did indeed play the same note continuously. But also, the same effect was heard as the band stood on the siding and the observer rolled past them on the flat car—first sharp, then true, then flat. What is important is not the velocity of either the observer or the source but rather their relative velocity. The amount of change in pitch is proportional to the relative velocity of the source and observer, divided by the speed of sound waves.

Light and other forms of electromagnetic radiation are also subject to the Doppler effect. In that case, the higher frequency of an approaching source corresponds to a bluer color of the observed wave, termed "blueshift," and a lower frequency ("redshift") for a receding source. You don't directly perceive the Doppler shift because the relevant wave velocity is the speed of light, and nothing on Earth goes fast enough for you to see the blue- or redshifts with your eyes. However, many people, perhaps including you, have been pulled over for speeding by police officers who used a hand-held combination radio transmitter and receiver ("hair dryer") that was sensitive enough to measure the tiny change in frequency of the radio echo off your car relative to the signal transmitted. (After all, you probably weren't driving at a substantial fraction of the speed of light.) Similarly, Doppler radar measurements reported TV programs indicate the speed of an approaching thunderstorm.

Virtual Astronomy Lab 3, "The Doppler Effect" extends your study of this phenomenon with further examples and calculations for sound waves in Section 2, then Doppler measurements of the motions of stars in Section 3. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

7

The Sun

Guidepost

The sun is the main source of light and warmth on Earth's surface, so it is a natural object of human curiosity and awe. It is also the star that is most easily visible from Earth. Knowledge of the interaction of light and matter, which you studied in Chapter 6, helps unlock the secrets of the sun and introduces you to the stars.

In this chapter, you will discover how analysis of the solar spectrum can paint a detailed picture of the sun's atmosphere and how basic physics has solved the mystery of what goes on in the sun's core. Here you will answer four important questions:

- **What can you learn about the sun by observing its surface and atmosphere?**
- **What are the dark sunspots?**
- **Why does the sun go through 11- and 22-year cycles of activity?**
- **What is the source of the sun's energy?**

Although this chapter considers only the star at the center of our solar system, introducing you to one star in detail leads you onward and outward in later chapters among the other stars that fill the universe.

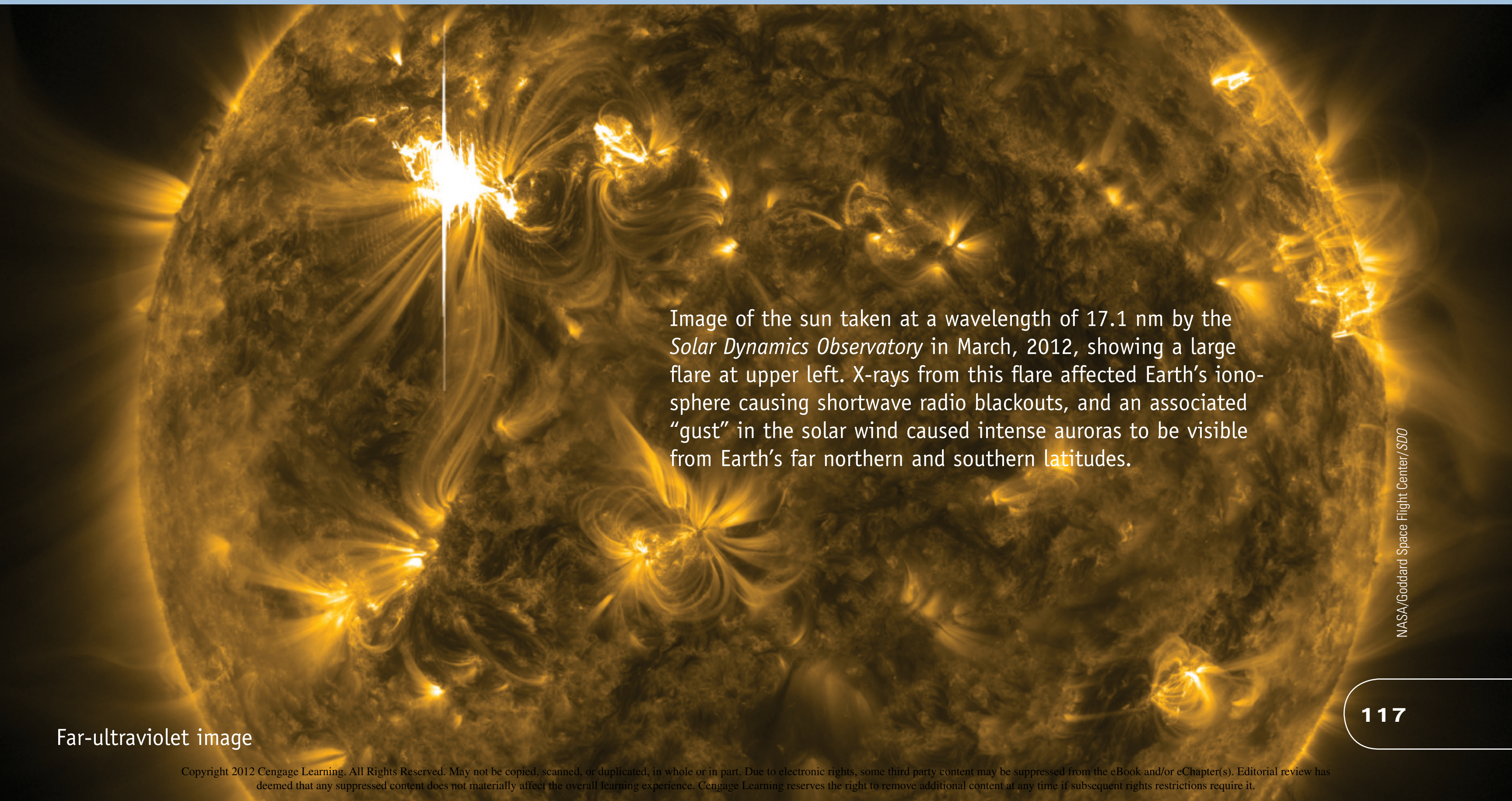


Image of the sun taken at a wavelength of 17.1 nm by the *Solar Dynamics Observatory* in March, 2012, showing a large flare at upper left. X-rays from this flare affected Earth's ionosphere causing shortwave radio blackouts, and an associated "gust" in the solar wind caused intense auroras to be visible from Earth's far northern and southern latitudes.

All cannot live on the piazza,
but everyone may enjoy the sun.

ITALIAN PROVERB

A SCIENTIST ONCE joked that solar astronomers would know a lot more about the sun if it were farther away. The sun is so close that Earth's astronomers can see swirling currents of gas and arched bridges of magnetic force at a level of detail that seems overwhelming. But, as you will learn in Chapter 13, the sun is just a normal star and, in a sense, it is a simple object. The sun is made up almost entirely of hydrogen and helium gas confined by its own gravity in a sphere 109 times Earth's diameter (**Celestial Profile 1**). The gases of the sun's atmosphere are hot and radiate the light and heat that make life possible on Earth. That part of the sun is where you can begin your exploration.

7-1 The Solar Atmosphere

THE SUN'S ATMOSPHERE IS MADE UP of three layers. The visible surface is the **photosphere,** and above that are the **chromosphere** and the **corona.** You first met these terms in Chapter 3 when you learned that you can distinguish these features during solar eclipses. [Note that astronomers normally speak of the interior of the sun as being "below" or "under" the photosphere, and the sun's atmosphere as being "above" or "over" the photosphere.]

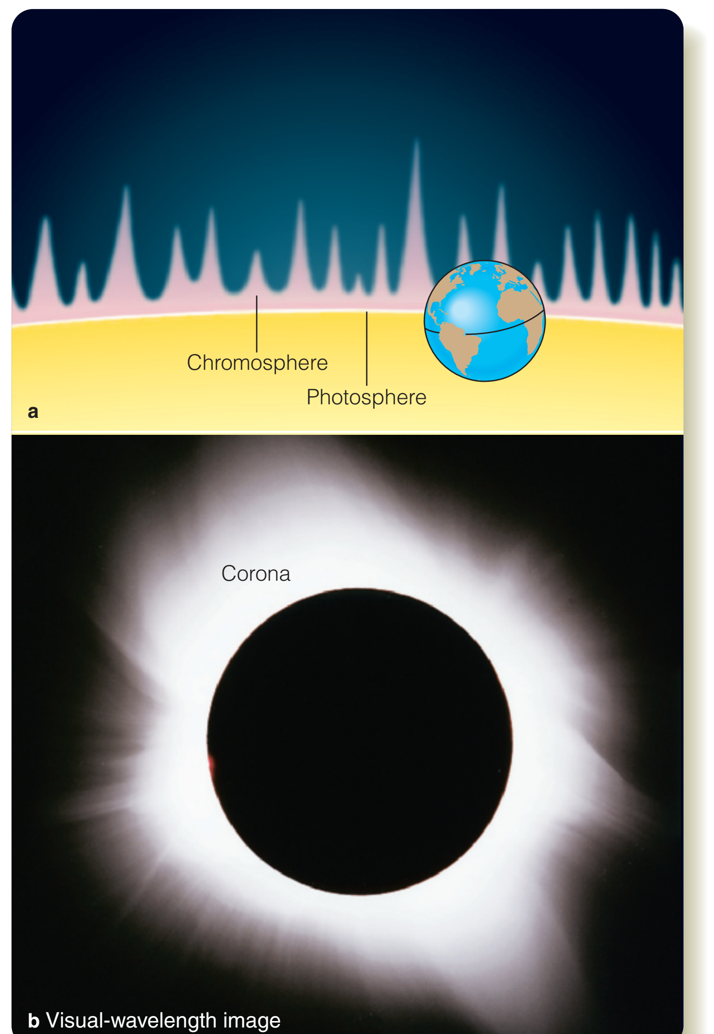
When you look at the sun you see the photosphere as a hot, glowing surface with a temperature of about 5800 K. That temperature is determined by precisely measuring the spectrum of sunlight, then using Wien's law (look back to Chapter 6). At that temperature, every square millimeter of the sun's surface is radiating more energy than a 60-watt light bulb (Stefan-Boltzmann law, Chapter 6). With all that energy radiating into space, the sun's surface would cool rapidly if energy did not flow up from the interior to keep the surface hot, so simple logic tells you that energy in the form of heat is flowing outward from the sun's interior. Not until the 1930s did astronomers understand that the sun makes its energy by nuclear reactions at the center. Those nuclear reactions are described in detail at the end of this chapter.

For now, you can consider the sun's atmosphere in its relatively quiet, average state. Later you can add details of its activity as heat flows up continuously from the sun's interior and makes the outer layers churn like a pot of boiling water.

The Photosphere

The visible surface of the sun looks like a smooth layer of gas marked only by a few dark **sunspots** that come and go over a few weeks. Although the photosphere seems to be a distinct surface, it is not solid. In fact, the sun is gaseous from its outer atmosphere right down to its center. The photosphere is the thin layer of gas from which Earth receives most of the sun's light. It is less than 500 km (300 mi) deep, and in a model of the sun the size of a bowling ball, the photosphere would be no thicker than a layer of tissue paper wrapped around the ball (■ Figure 7-1).

The photosphere is the layer in the sun's atmosphere that is dense enough to emit plenty of light but not so dense that light can't escape. Below the photosphere, the gas is denser and hotter and therefore radiates plenty of light, but that light cannot escape because it is blocked by the outer layers of gas. Therefore, Earth does not receive light from deeper layers under the photosphere. In contrast, the gas above the photosphere is less dense, so although Earth does receive that light, there is not much of it.



■ Figure 7-1

(a) A cross section at the edge of the sun shows the relative thickness of the photosphere and chromosphere. Earth is shown for scale. On this scale, the disk of the sun would be more than 1.5 m (5 ft) in diameter. The corona extends from the top of the chromosphere to great height above the photosphere. (b) This photograph, made during a total solar eclipse, shows only the inner part of the corona.

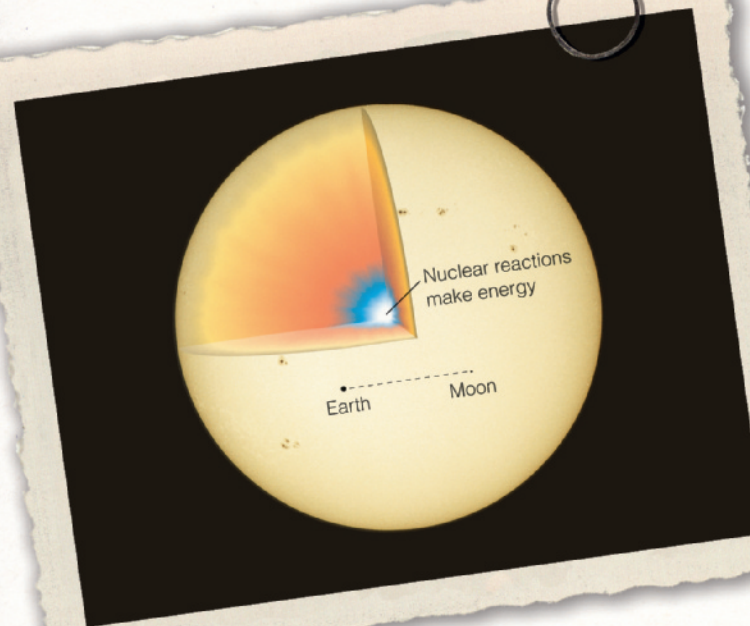
The photosphere appears to be substantial, but it is really a very low-density gas. Even in its deepest and densest layers, the photosphere is less than 1/3000 as dense as the air you breathe. To find gases as dense as the air at Earth's surface, you would have to descend about 15,000 km (10,000 mi) below the photosphere. With fantastically efficient insulation, you could fly a spaceship right through the photosphere.

The spectrum of the sun is an absorption spectrum, and that can tell you a great deal about the photosphere. You know from Kirchhoff's third law (look back again to Chapter 6) that an absorption spectrum is produced when the source of a continuous (blackbody) spectrum is viewed through a transparent gas. The deeper layers of the photosphere are dense enough to produce a continuous spectrum. Atoms in higher, transparent layers of the photosphere and in the sun's atmosphere absorb photons with the unique wavelengths and energies corresponding to the jumps between each type of atom's electron orbits, producing absorption lines that allow you to identify hydrogen, helium, and other elements.

In good photographs, the photosphere has a mottled appearance because it is made up of dark-edged regions called granules. The overall pattern is called **granulation** (■ Figure 7-2a). Granules can be several thousand kilometers across (as large as the state of Texas) but last for only 10 to 20 minutes each before fading, shrinking, and being replaced by a new granule. Detailed observations of the granules show that their centers emit more blackbody radiation and are slightly bluer than the edges. From this information plus the Wien and Stefan-Boltzmann laws, astronomers can calculate that the granule centers are a few hundred degrees hotter than the edges (Figure 7-2b). Doppler shifts of spectral lines (Chapter 6) reveal that the centers are rising and the edges are sinking at speeds of about 0.4 km/s (900 mph).

From this evidence, astronomers recognize granulation as the surface effects of convection currents just below the photosphere. **Convection** occurs when hot material rises and cool material sinks, as when, for example, a current of hot gas rises above a candle flame. You can watch convection in a liquid by adding a bit of cool nondairy creamer to an unstirred cup of hot coffee. The cool creamer sinks, gets warmer, expands, rises, cools, contracts, sinks again, and so on, creating small regions on the surface of the coffee that mark the tops of convection currents. Viewed from above, these coffee regions look something like solar granules. The presence of granulation is clear evidence that energy is flowing upward through the photosphere. You will learn more about the sun's convection currents and internal structure later in this chapter.

Spectroscopic studies of the solar surface have revealed another less obvious kind of granulation. **Supergranules** are regions a little over twice Earth's diameter that include about 300 granules each. These supergranules are regions of very slowly rising currents that last a day or two. They appear to be produced by larger gas currents that lie deeper under the photosphere than the ones that produce the granules.



© Cengage Learning 2014; image: Daniel Good

This visible image of the sun shows a few sunspots and is cut away to show the location of energy generation at the sun's center. The Earth–moon system is shown for scale.

Celestial Profile 1: The Sun

From Earth:

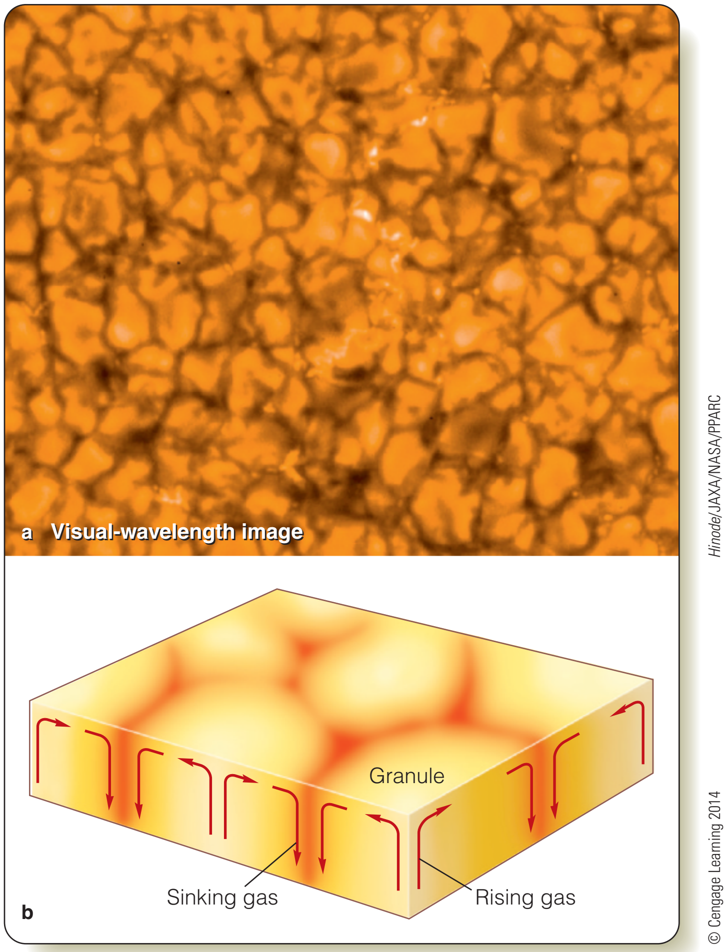
Average distance from Earth	1.000 AU (1.496×10^8 km)
Maximum distance from Earth	1.017 AU (1.521×10^8 km)
Minimum distance from Earth	0.983 AU (1.471×10^8 km)
Equatorial angular diameter	0.533° (1920 arc seconds)
Period of rotation (sidereal)	24.5 days at equator
Apparent visual magnitude	−26.74

Characteristics:

Radius	6.96×10^5 km
Mass	1.99×10^{30} kg
Average density	1.41 g/cm^3
Escape velocity at surface	618 km/s
Luminosity	3.84×10^{26} J/s
Surface temperature	5780 K
Central temperature	15.7×10^6 K
Spectral type	G2 V
Absolute visual magnitude	+4.83

Personality Profile:

In Greek mythology, the sun was carried across the sky in a golden chariot pulled by powerful horses and guided by the sun god Helios. When Phaeton, the son of Helios, drove the chariot one day, he lost control of the horses, and Earth was nearly set ablaze before Zeus smote Phaeton from the sky. Even in classical times, people understood that life on Earth depends critically on the sun.



■ **Figure 7-2**

(a) This ultra-high-resolution image of the photosphere shows granulation. The largest granules here are about the size of Texas. (b) This model explains granulation as the tops of rising convection currents just below the photosphere. Heat flows upward as rising currents of hot gas and downward as sinking currents of cool gas. The rising currents heat the solar surface in small regions seen from Earth as granules.

The edge, or limb, of the solar disk is dimmer than the center (see the figure in Celestial Profile 1). This limb darkening is caused by the absorption of light in the photosphere. When you look at the center of the solar disk, you are looking directly down into the sun, and you see deep, hot, bright layers in the photosphere. In contrast, when you look near the limb of the solar disk, you are looking at a steep angle to the surface and cannot see as deeply. The photons you see come from shallower, cooler, dimmer layers in the photosphere. Limb darkening proves that the temperature in the photosphere increases with depth, yet another confirmation that energy is flowing up from below.

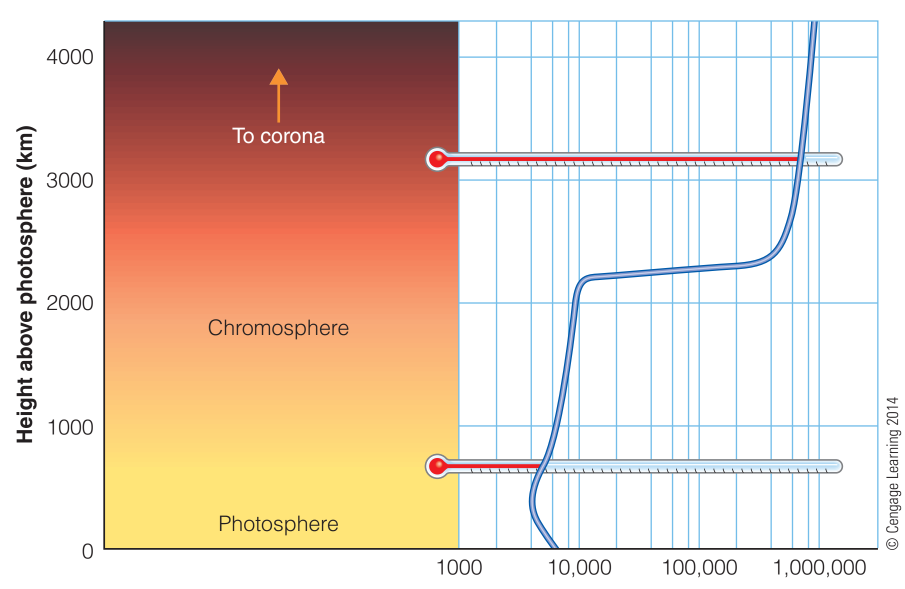
The Chromosphere

Above the photosphere lies the chromosphere. Solar astronomers define the lower edge of the chromosphere as lying just above the visible surface of the sun, with its upper regions blending gradually with the corona. You can think of the chromosphere as an irregular layer with a depth on average less than Earth's diameter (see Figure 7-1a). Because the chromosphere is roughly 1000 times fainter than the photosphere, you can see it with your unaided eyes only during a total solar eclipse when the moon covers the brilliant photosphere. Then, the chromosphere flashes into view as a thin pink layer just above the photosphere. The word *chromosphere* comes from the Greek word *chroma,* meaning "color." The pink color is produced by the combined light of three bright emission lines—the red, blue, and violet Balmer lines of hydrogen (look back once more to Chapter 6).

The chromosphere produces an emission spectrum, and Kirchhoff's second law tells you it therefore must be an excited, low-density gas. The chromosphere's density ranges from ten thousand (at the bottom, near the photosphere) to one hundred billion (at the top, near the corona) times less dense than the air you breathe. Further analysis of solar spectra reveals that atoms in the lower chromosphere are ionized, and atoms in the higher layers of the chromosphere are even more highly ionized. That is, they have lost more electrons. Hydrogen atoms contain only one electron, but atoms like calcium and iron contain many more, and at higher temperatures they can lose several electrons and become highly ionized. Astronomers can find the temperature in different parts of the chromosphere from the amount of ionization. Just above the photosphere the temperature falls to a minimum of about 4500 K and then rises rapidly (■ Figure 7-3) to the extremely high tem-

■ **Figure 7-3**

The chromosphere. If you could place thermometers in the sun's atmosphere, you would discover that the temperature increases from 5800 K at the photosphere to 1,000,000 K at the top of the chromosphere.



peratures of the corona. The upper chromosphere is hot enough to emit X-rays and can be studied by X-ray telescopes in space.

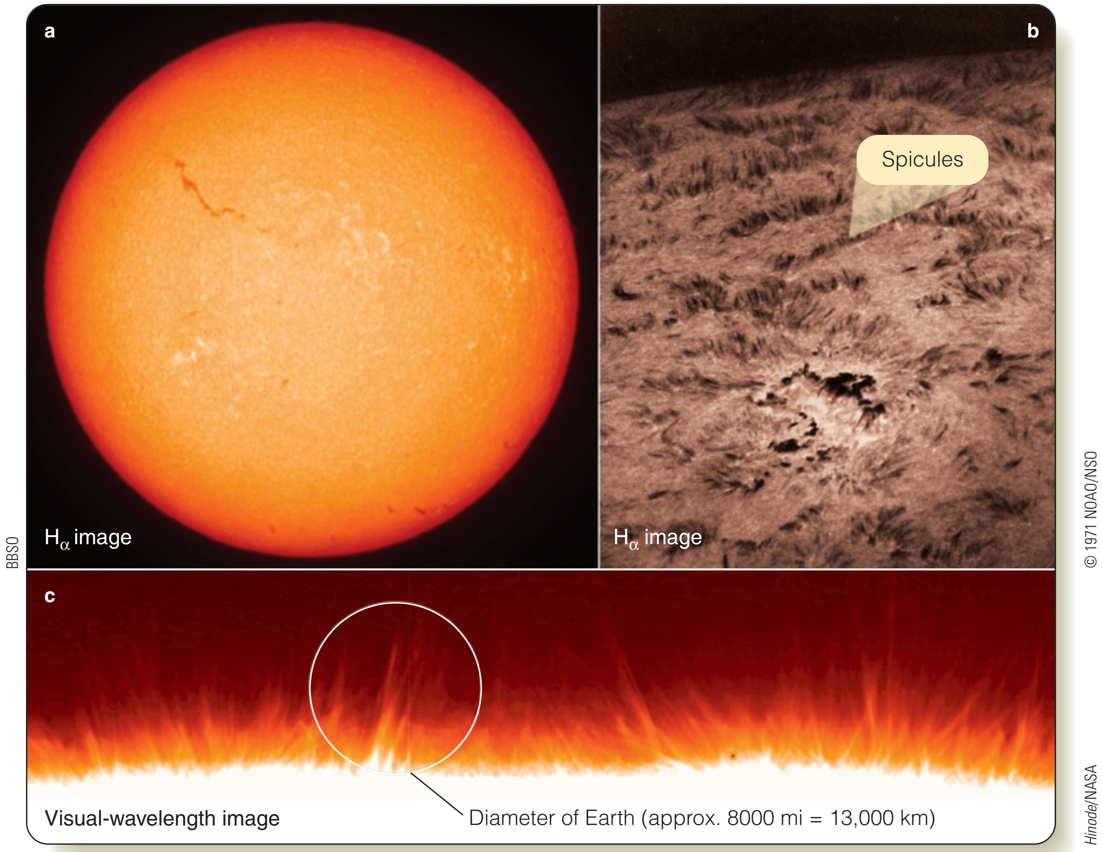
Solar astronomers can take advantage of the way spectral lines form to map the chromosphere. The gases of the chromosphere are transparent to nearly all visible light, but atoms in that gas are very good at absorbing photons of a few specific wavelengths. This produces some exceptionally strong (dark) absorption lines in the solar spectrum. A photon at one of those wavelengths is unlikely to escape from deeper layers to be received at Earth and can only come to you from higher in the sun's atmosphere. A **filtergram** is an image of the sun made using light only at the wavelength of one of those strong absorption lines to reveal detail in the upper regions of the chromosphere. Another way to study these layers of gas high in the sun's atmosphere is to record solar images in the far-ultraviolet or X-ray portions of the electromagnetic spectrum, because those layers are very hot and emit most of their light at short wavelengths.

■ Figure 7-4 shows a filtergram made at the wavelength of the H_α Balmer line. This image shows complex structure in the chromosphere. **Spicules** are flamelike jets of gas extending upward into the chromosphere and lasting 5 to 15 minutes. Seen at the limb of the sun's disk, these spicules blend together and look like flames covering a burning prairie (Figure 7-4c), but they are more like the opposite of flames; spectra show that spicules are cooler gas from the lower chromosphere extending upward into hotter regions. Images at the center of the solar disk show that spicules spring up around the edge of supergranules like weeds around flagstones (Figure 7-4b).

The Solar Corona

The outermost part of the sun's atmosphere is called the corona, after the Greek word for crown. The corona is so dim that, like the chromosphere, it is not visible in Earth's daytime sky because of the glare of scattered light from the sun's brilliant photosphere. During a total solar eclipse, the innermost parts of the corona are visible to the unaided eye, as shown in Figure 7-1b (also, look back to Chapter 3). Observations made with specialized telescopes called **coronagraphs** can block the light of the photosphere and record the corona out beyond 20 solar radii, almost 10 percent of the way to Earth. Such images reveal streamers in the corona that follow magnetic lines of force in the sun's magnetic field (■ Figure 7-5). Later in this chapter you will learn more about how features and activity in the sun's atmosphere are controlled by magnetic fields.

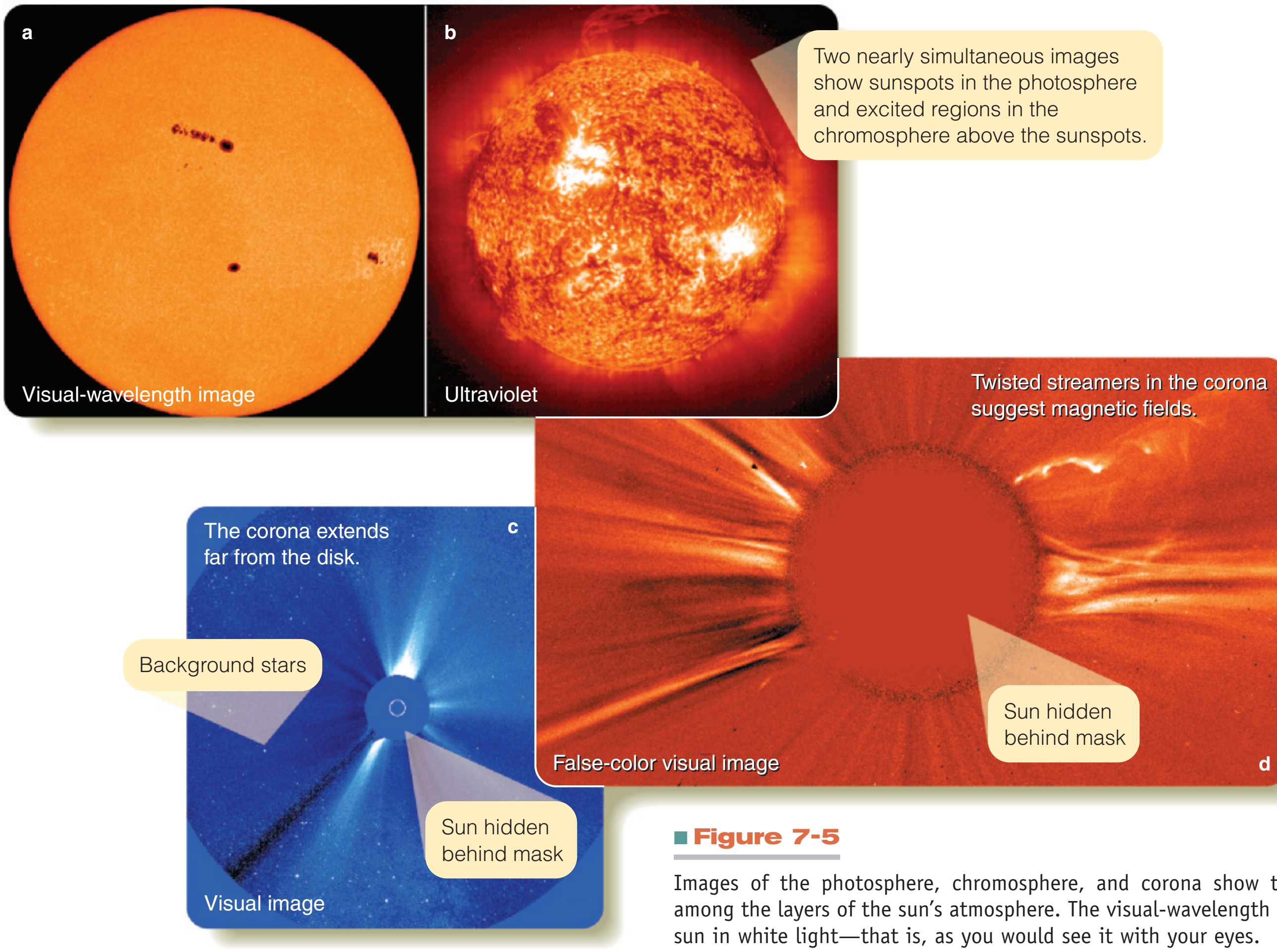
The corona's spectrum, like that of the upper chromosphere, includes emission lines of highly ionized gases. In the lower corona, the atoms are not as highly ionized as they are at higher altitudes, and this tells you that the temperature of the corona rises with altitude. Just above the chromosphere, the temperature is about 500,000 K, but in the outer corona the temperature can be 2 million K or more. The corona is hot enough to emit X-rays, but the coronal gas is not very bright because its density is low, only 10^6 atoms/cm^3 in its lower regions. That is about 10 billion times less dense than the air you breathe. In its outer layers the corona contains only about 10^5 atoms/cm^3, fewer than in the best vacuum in laboratories on Earth.



© Cengage Learning 2014

■ Figure 7-4

H_α filtergrams reveal complex structure in the chromosphere that cannot be seen at visual wavelengths, including spicules springing from the edges of supergranules over twice the diameter of Earth. Seen at the edge of the solar disk, spicules look like a burning prairie, but they are not at all related to burning. Compare with Figure 7-1a.



© Cengage Learning 2014; all images: *SOHO*/ESA/NASA

Figure 7-5

Images of the photosphere, chromosphere, and corona show the relationships among the layers of the sun's atmosphere. The visual-wavelength image shows the sun in white light—that is, as you would see it with your eyes.

Astronomers have wondered for years how the corona and chromosphere can be so hot. Heat flows from hot regions to cool regions, never from cool to hot. So how can the heat from the photosphere, with a temperature of only 5800 K, flow out into the much hotter chromosphere and corona? Observations made by the *SOHO (Solar and Heliospheric Observatory)* satellite have mapped a **magnetic carpet** of looped magnetic fields extending up through the photosphere. Because the gas in the chromosphere and the corona has very low densities, it can't resist movements of the magnetic fields. Turbulence below the photosphere seems to flick the magnetic loops back and forth and whip the gas about, heating the gas. Furthermore, observations with the *Hinode* (pronounced *Hee-noh-day*) spacecraft have revealed magnetic waves generated by turbulence below the photosphere traveling up into the chromosphere and corona and heating the gas. In both cases, energy appears to flow out from the interior of the sun to the chromosphere and corona not by radiation but by agitation of magnetic fields.

Ionized, low-density gas cannot cross magnetic fields, so where the sun's field loops back to the surface, the gas is trapped. However, some of the magnetic fields lead outward into space, and there the gas flows away from the sun in the **solar wind** that can be considered an extension of the corona. The low-density gases of the solar wind blow past Earth at 300 to 800 km/s with gusts as high as 1000 km/s (more than 2 million mph). Earth is bathed in the corona's hot breeze.

Because of the solar wind, the sun is slowly losing mass, but this is only a minor loss for an object as massive as the sun. The sun loses about 10 million tons per second, but that is only 10^{-14} of its mass per year. Later in life, the sun, like many other stars, will lose mass rapidly in a more powerful wind. You will see in future chapters how rapid outflowing winds affect the evolution of stars.

Do other stars have chromospheres, coronae (plural of corona), and stellar winds like the sun? Stars are so far away they never look like more than points of light, but ultraviolet and X-ray observations suggest that the answer is yes. The spectra of many stars contain emission lines at far-ultraviolet wavelengths that could have formed only in the low-density, high-temperature gases of a chromosphere and corona. Also, many stars are sources of X-rays, which appear to have been produced by high-temperature gas in their chromospheres and coronae. This observational evidence gives astronomers good reason to consider the sun, despite all its complexity seen from our nearby viewpoint, to be a typical star.

Composition of the Sun

It seems as though it should be easy to learn the composition of the sun just by studying its spectrum, but this is actually a difficult problem that wasn't well understood until the 1920s. The solution of that problem is the story of an important American astronomer who waited decades to get proper credit for her work.

In her 1925 Ph.D. thesis work, Cecilia Payne invented modern methods of interpreting spectra of the sun and stars. For example, sodium lines are observed in the sun's spectrum, so you can be sure that the sun's atmosphere contains some sodium atoms. Payne came up with a mathematical procedure to determine just how many sodium atoms are there. She also proved that if spectral lines of a certain element are not detected in the sun's spectrum, that element might still be present, but the gas is too hot or too cool or the wrong density for that type of atom to have electrons in the right energy levels to produce visible lines.

Payne's first calculations showed that over 90 percent of the atoms in the sun must be hydrogen and most of the rest are helium (■ Table 7-1). In contrast, atoms like calcium, sodium, and iron that have strong lines in the sun's spectrum are actually not very abundant. Rather, at the temperature of the sun's atmosphere, those atoms are especially efficient at absorbing photons with wavelengths of visible light.

At the time she did her original work, astronomers found it hard to believe Payne's calculated abundances of hydrogen, helium, and other elements in the sun. They especially found such a high abundance of helium unacceptable because helium lines are nearly invisible in the sun's spectrum. Eminent astronomers dismissed Payne's result as illusory; it was only several decades later that the scientific community realized the value of her work. Now we know that she was completely correct. Payne's work on the composition of the sun illustrates the importance of fully understanding the interaction between light and matter in order to understand objects in the universe. In Chapter 13 you will continue the story of Cecilia Payne with regard to the main part of her thesis, interpreting the spectra of stars other than the sun.

The layers of the solar atmosphere are all that astronomers can observe directly, but there are phenomena in those layers that reveal what it's like inside the sun—your next destination.

■ Table 7-1 | The Most Abundant Elements in the Sun

Element	Percentage by Number of Atoms	Percentage by Mass
Hydrogen	91.0	70.6
Helium	8.9	27.5
Carbon	0.03	0.3
Nitrogen	0.01	0.1
Oxygen	0.08	1.0
Neon	0.01	0.2
Magnesium	0.003	0.05
Silicon	0.003	0.07
Sulfur	0.002	0.04
Iron	0.003	0.1

Below the Photosphere

Almost no light emerges from below the photosphere, so you can't see into the solar interior. However, solar astronomers using a technique called **helioseismology** can analyze naturally occurring vibrations in the sun to explore its depths. Convective movements of gas in the sun constantly produce vibrations—rumbles that would be much too low to hear with human ears even if your ears could survive a visit to the sun's atmosphere. Some of these vibrations resonate in the sun like sound waves in organ pipes. A vibration with a period of 5 minutes is strongest, but the periods range from 3 to 20 minutes. These are very, very low-pitched sounds!

Astronomers can detect these vibrations by observing Doppler shifts in the solar surface. As a vibrational wave travels down into the sun's interior, the changing density and temperature of the gas it is moving through curves its path, and it returns to the surface, where it makes the photosphere heave up and down by small amounts—roughly plus or minus 15 km (10 mi). Multiple vibrations occurring simultaneously cover the surface of the sun with a pattern of rising and falling regions that can be mapped using the Doppler effect (■ Figure 7-6). By observing these motions, astronomers can determine which vibrations resonate and become stronger, versus which become weaker. Short-wavelength waves penetrate less deeply and travel shorter distances than longer-wavelength waves, so the vibrations of different wavelengths explore different layers in the sun. Just as geologists can study Earth's interior by analyzing vibrations from earthquakes, so solar astronomers can use helioseismology to explore the sun's interior.

Helioseismology requires huge amounts of data, so astronomers have used a network of telescopes around the world operated by the Global Oscillation Network Group (GONG). The network can observe the sun continuously for weeks at a time as Earth rotates; in other words, the sun never sets on GONG. The *SOHO* satellite in space can observe solar oscillations continuously and can detect motions as slow as 1 mm/s (0.002 mph). Solar astronomers can then use supercomputers to distinguish the different vibration patterns on the solar surface and measure the strength of the waves at many different wavelengths.

Helioseismology has allowed astronomers to map the temperature, density, and rate of rotation in the interior of the sun, as well as to find the positions and speeds of great currents of gas flowing below the photosphere. That information confirms a model developed to understand the cycle of sunspots and solar activity that you will learn about in the next section.

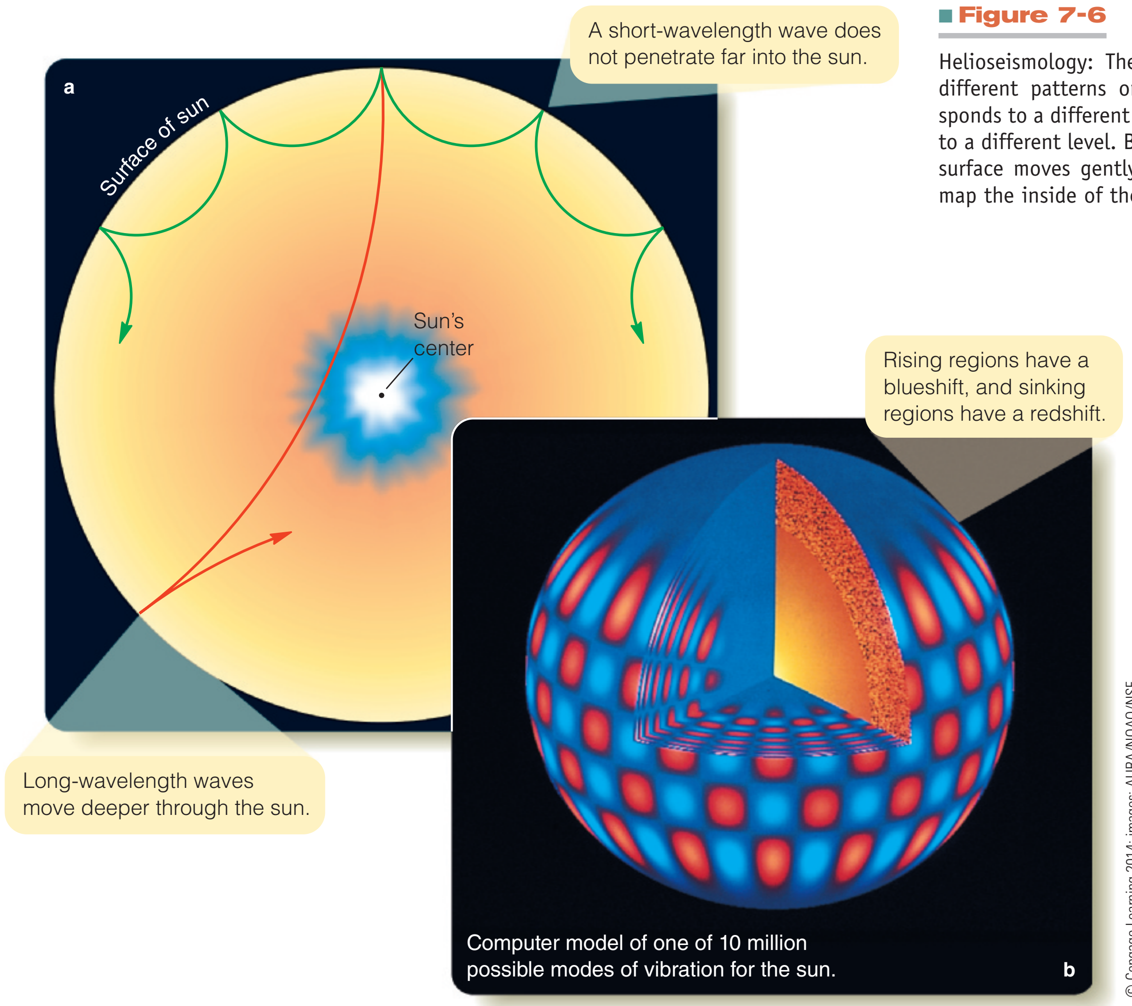


Figure 7-6

Helioseismology: The sun can vibrate in millions of different patterns or modes, and each mode corresponds to a different wavelength vibration penetrating to a different level. By measuring Doppler shifts as the surface moves gently up and down, astronomers can map the inside of the sun.

SCIENTIFIC ARGUMENT

What evidence leads astronomers to conclude that temperature increases with height in the chromosphere and corona?

Scientific arguments usually involve evidence, and in astronomy evidence means observations. Solar astronomers can observe the spectrum of the chromosphere, and they find that atoms there are more highly ionized (have lost more electrons) than atoms in the photosphere. Atoms in the corona are even more highly ionized. That must mean the chromosphere and corona are hotter than the photosphere.

Evidence is the key to understanding how science works. Now it is time to build a new argument. **What evidence leads astronomers to conclude that other stars have chromospheres and coronae like those of the sun?**

7-2 Solar Activity

THE SUN IS NOT QUIET. It has slowly changing storms larger than Earth that we see as sunspots and unimaginably vast eruptions. All of these seemingly different forms of solar activity have one thing in common—magnetic fields. The weather on the sun is magnetic.

Observing the Sun

Solar activity is often visible with even a small telescope, but you should be very careful if you try to observe the sun. Sunlight is intense, and when it enters your eye it is absorbed and converted into thermal energy. The infrared radiation in sunlight is especially dangerous because your eyes can't detect it. You don't sense how intense the infrared is, but it is converted to thermal energy in your eyes and can burn and scar your retinas.

It is not safe to look directly at the sun, and it is even more dangerous to look at the sun through any optical instrument such as a telescope, binoculars, or even the viewfinder of a camera. The light-gathering power of such an optical system concentrates the sunlight and can cause severe injury. Never look at the sun with any optical instrument unless you are certain it is safe. ■ Figure 7-7a shows a safe way to observe the sun with a small telescope.

In the early 17th century, Galileo observed the sun with a thick dark filter over his telescope and saw spots on its surface. Day by day, he saw the spots moving across the sun's disk and rightly

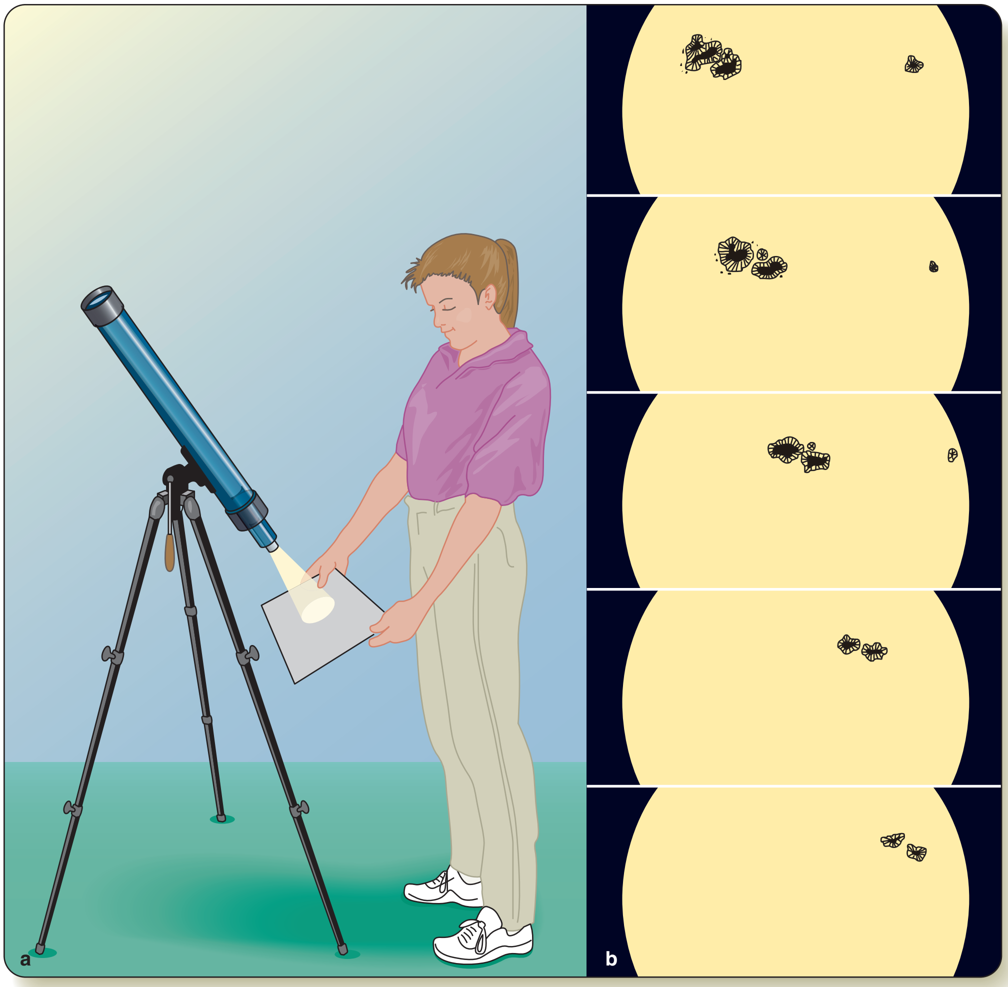


Figure 7-7

(a) Looking through a telescope at the sun is dangerous, but you can always view the sun safely with a small telescope by projecting its image on a white screen. (b) If you sketch the location and structure of sunspots on successive days, you will see the rotation of the sun and gradual changes in the size and structure of sunspots, just as Galileo did in 1610.

concluded that the sun is a rotating sphere. If you repeated Galileo's observations, you would probably also detect sunspots, a view that would look something like Figure 7-7b.

Sunspots

The dark sunspots that you see at visible wavelengths only hint at the complex processes that go on in the sun's atmosphere. To explore those processes, you need to analyze images and spectra at a wide range of wavelengths.

Study **Sunspots and the Sunspot Cycle** on pages 126–127 and notice five important points and four new terms:

1. Sunspots are cool spots on the sun's photosphere, usually appearing in groups, that form and disappear over time scales of weeks and months.
2. Sunspot numbers follow an 11-year cycle, becoming more numerous, reaching a maximum, and then becoming much less numerous. The *Maunder butterfly diagram* shows how the location of sunspots also changes during a cycle.
3. The *Zeeman effect* gives astronomers a way to measure the strength of magnetic fields on the sun and provide evidence that sunspots contain, and are caused by, locally strong magnetic fields. When the magnetic properties of sunspots are considered, the 11-year cycle is found to be really a 22-year cycle.
4. The intensity of the sunspot cycle can vary from cycle to cycle and appears to have almost faded away during the *Maunder minimum* in the late 17th century. That seems to have affected Earth's climate.
5. The evidence is clear that sunspots are parts of *active regions* dominated by magnetic fields that involve all layers of the sun's atmosphere.

The sunspot groups are merely the visible traces of magnetically active regions. But what causes this magnetic activity? The answer is linked to the growth and decay cycle of the sun's overall magnetic field.

The Sun's Magnetic Cycle

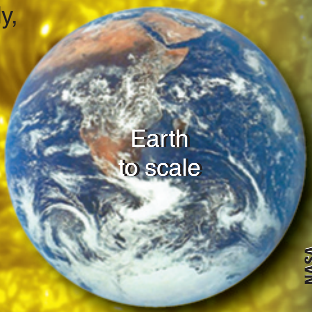
You are probably familiar with magnetic fields from classroom demonstrations with magnets and iron filings and from seeing the effect of Earth's magnetic field on a compass needle. The sun's

Sunspots and the Sunspot Cycle

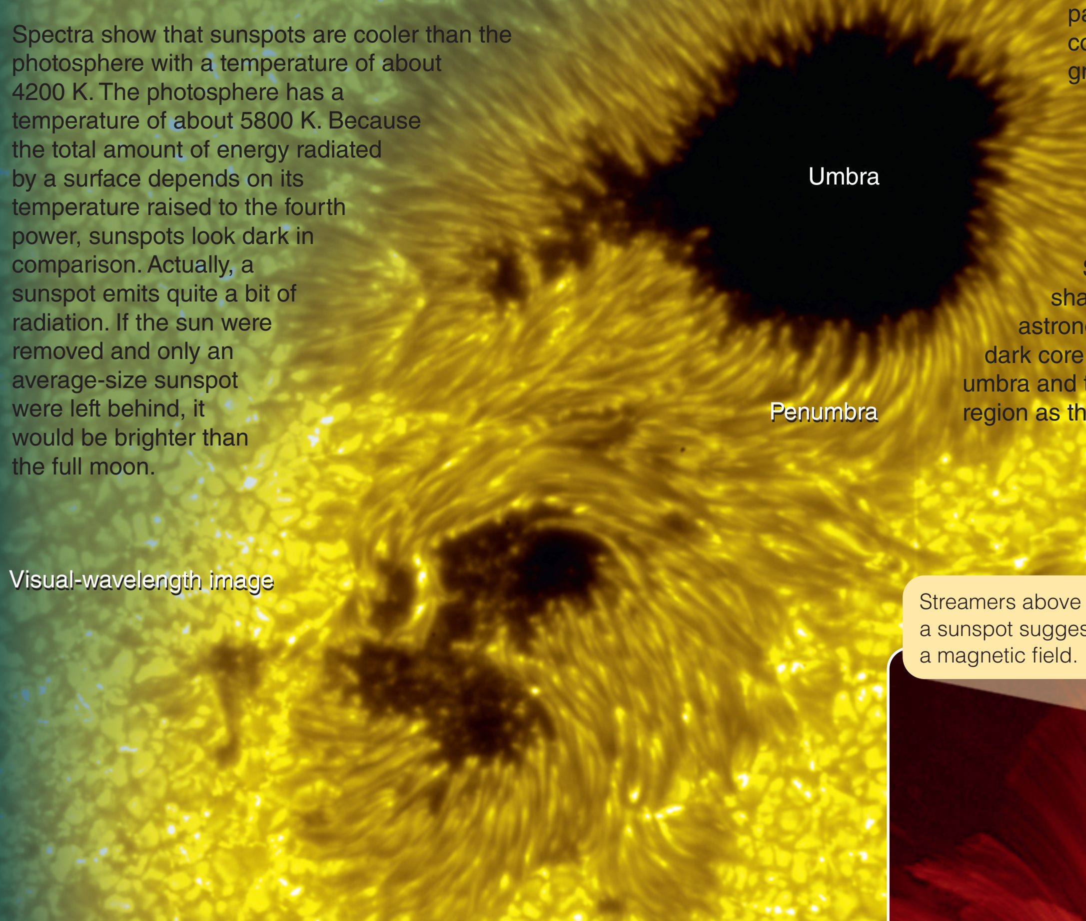
1 The dark spots that appear on the sun are only the visible traces of complex regions of activity. Observations over many years and at a range of wavelengths tell you that sunspots are clearly linked to the sun's magnetic field.

Spectra show that sunspots are cooler than the photosphere with a temperature of about 4200 K. The photosphere has a temperature of about 5800 K. Because the total amount of energy radiated by a surface depends on its temperature raised to the fourth power, sunspots look dark in comparison. Actually, a sunspot emits quite a bit of radiation. If the sun were removed and only an average-size sunspot were left behind, it would be brighter than the full moon.

A typical sunspot is about twice the size of Earth, but there is a wide range of sizes. They appear, last a few weeks to as long as 2 months, and then shrink away. Usually, sunspots occur in pairs or complex groups.



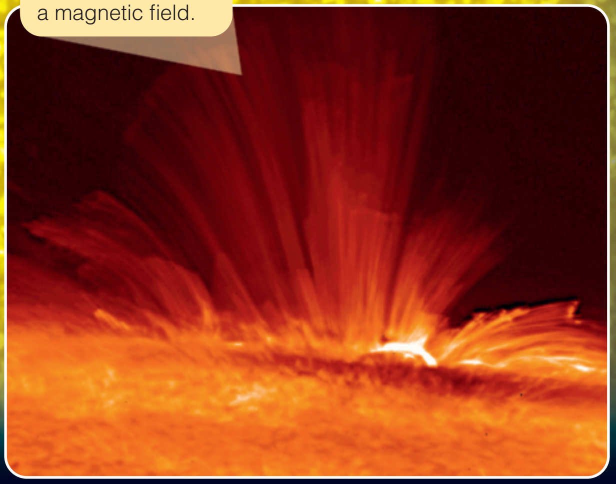
NASA



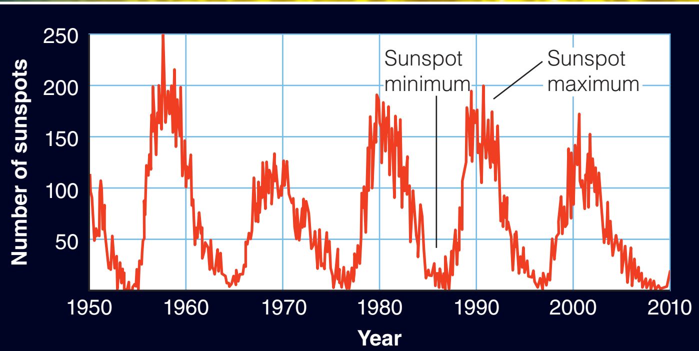
Sunspots are not shadows, but astronomers refer to the dark core of a sunspot as its umbra and the outer, lighter region as the penumbra.

Hinode JAXA/NASA

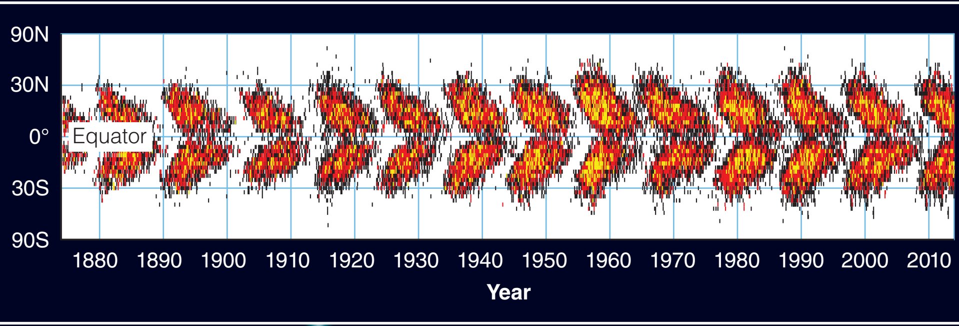
Streamers above a sunspot suggest a magnetic field.



Hinode JAXA/NASA



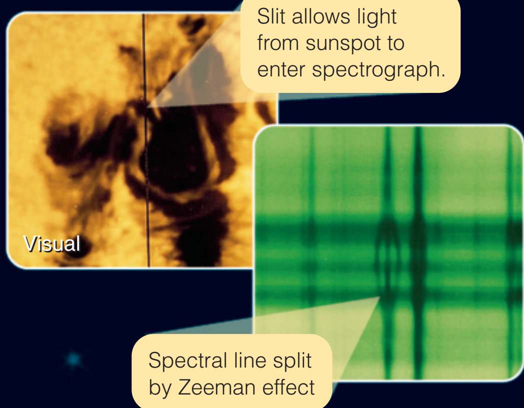
2 The number of spots visible on the sun varies in a cycle with a period of 11 years. At maximum, there are often over 100 spots visible. At minimum, there are very few.

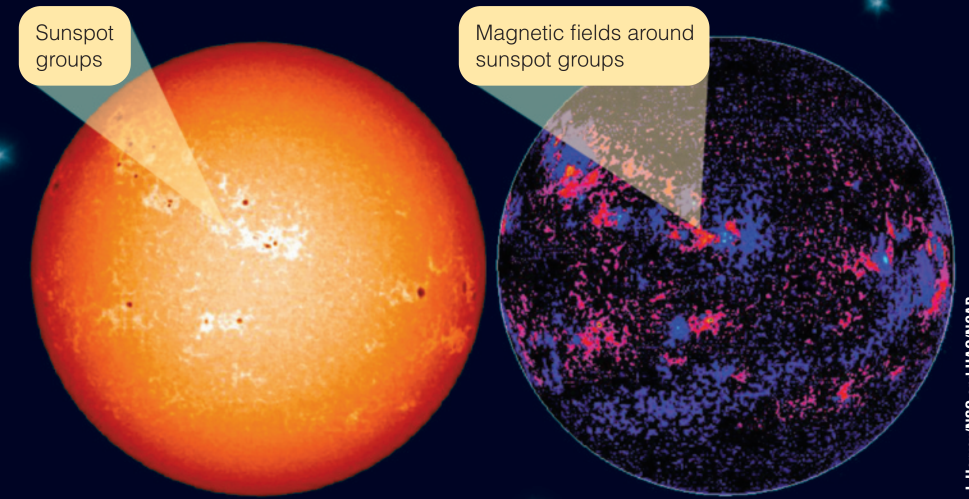


D. Hathaway NASA/MSFC

2a Early in the cycle, spots appear at high latitudes north and south of the sun's equator. Later in the cycle, the spots appear closer to the sun's equator. If you plot the latitude of sunspots versus time, the graph looks like butterfly wings, as shown in this **Maunder butterfly diagram,** named after E. Walter Maunder of Greenwich Observatory.

3 Astronomers can measure magnetic fields on the sun using the **Zeeman effect** as shown below. When an atom is in a magnetic field, the electron orbits are altered, and the atom is able to absorb a number of different wavelength photons even though it was originally limited to a single wavelength. In the spectrum, you see single lines split into multiple components, with the separation between the components proportional to the strength of the magnetic field.



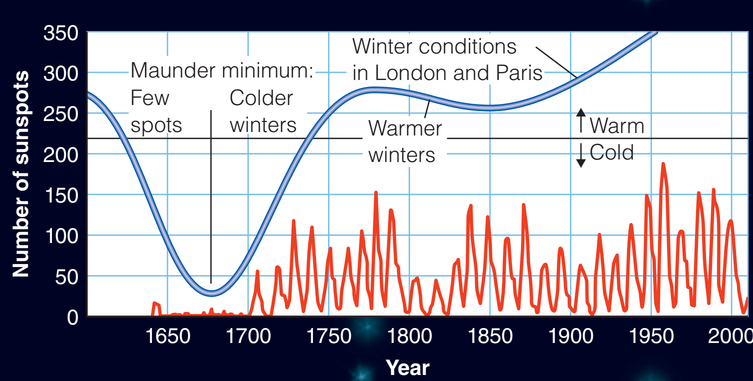


Ultraviolet filtergram

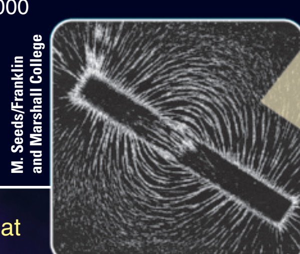
Magnetic image

Simultaneous images

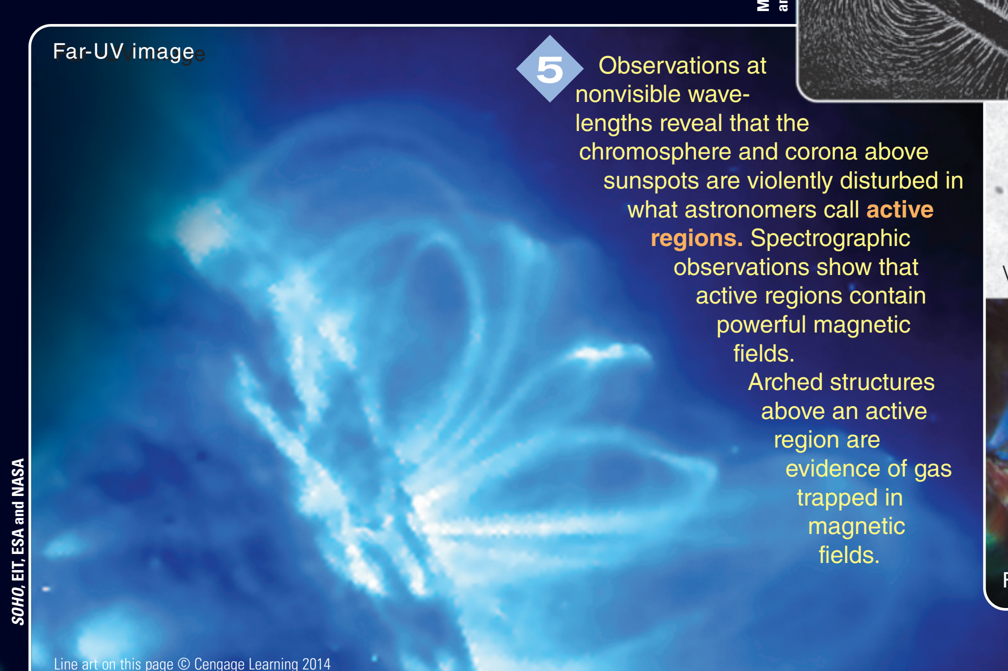
3a Images of the sun above show that sunspots contain magnetic fields a few thousand times stronger than Earth's. The strong fields are thought to inhibit gas motion below the photosphere; consequently, convection is reduced below the sunspot, and the surface there is cooler. Heat prevented from emerging through the sunspot is deflected and emerges around the sunspot, which can be detected in ultraviolet and infrared images.



4 Historical records show that there were very few sunspots from about 1645 to 1715, a phenomenon known as the **Maunder minimum.** This coincides with the middle of a period called the "Little Ice Age," a time of unusually cool weather in Europe and North America from about 1500 to about 1850, as shown in the graph at left. Other such periods of cooler climate are known. The evidence suggests that there is a link between solar activity and the amount of solar energy Earth receives. This link has been confirmed by measurements made by spacecraft above Earth's atmosphere.

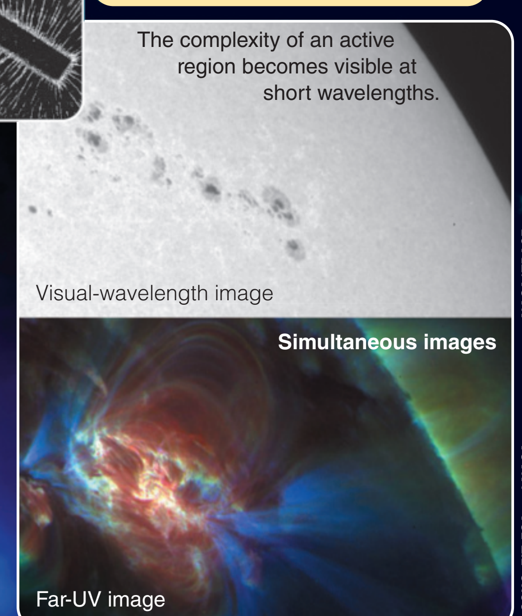


Magnetic fields can reveal themselves by their shape. For example, iron filings sprinkled over a bar magnet reveal an arched shape.



5 Observations at nonvisible wavelengths reveal that the chromosphere and corona above sunspots are violently disturbed in what astronomers call **active regions.** Spectrographic observations show that active regions contain powerful magnetic fields.

Arched structures above an active region are evidence of gas trapped in magnetic fields.



Line art on this page © Cengage Learning 2014

magnetic field is powered by the energy flowing outward through the moving currents of gas. The gas is highly ionized, so it is a very good conductor of electricity. When that electrically conducting gas rotates and is stirred by convection, some of the energy of rotation and also energy flowing outward by convection can be converted into a magnetic field. This process is called the **dynamo effect,** and it is understood also to operate in Earth's core and produce Earth's magnetic field. Helioseismologists have found evidence that the dynamo effect generates the sun's magnetic field at the bottom of the convection currents, deep under the photosphere.

The sun's magnetic field cannot be as stable as Earth's. The sun does not rotate as a rigid body. It is a gas from its outermost layers down to its center, so some parts of the sun can rotate faster than other parts. The equatorial region of the photosphere rotates faster than do regions at higher latitudes (■ Figure 7-8a). At the equator, the photosphere rotates once every 24.5 days, but at latitude 45° one rotation takes 27.8 days. This phenomenon is called **differential rotation.** Helioseismology maps of the rotation in the sun's interior (Figure 7-8b) reveal that the gas at different levels also rotate with different periods, another type of differential rotation. Both types of differential rotation seem to be involved in the sun's magnetic cycle.

Although the magnetic cycle is not fully understood, the **Babcock model** (named for the scientist who invented it) explains the magnetic cycle as repeated tangling and untangling of the solar magnetic field. Because an ionized gas is a very good conductor of electricity, if the gas moves, the embedded electrical currents and resulting magnetic fields must move with it. As a result, differential rotation drags the magnetic field along and wraps it around the sun like a long string caught on a turning wheel. Rising and sinking convection currents then twist the field into ropelike tubes. The model predicts that pairs of sunspots should occur where these magnetic tubes burst through the sun's surface (■ Figure 7-9).

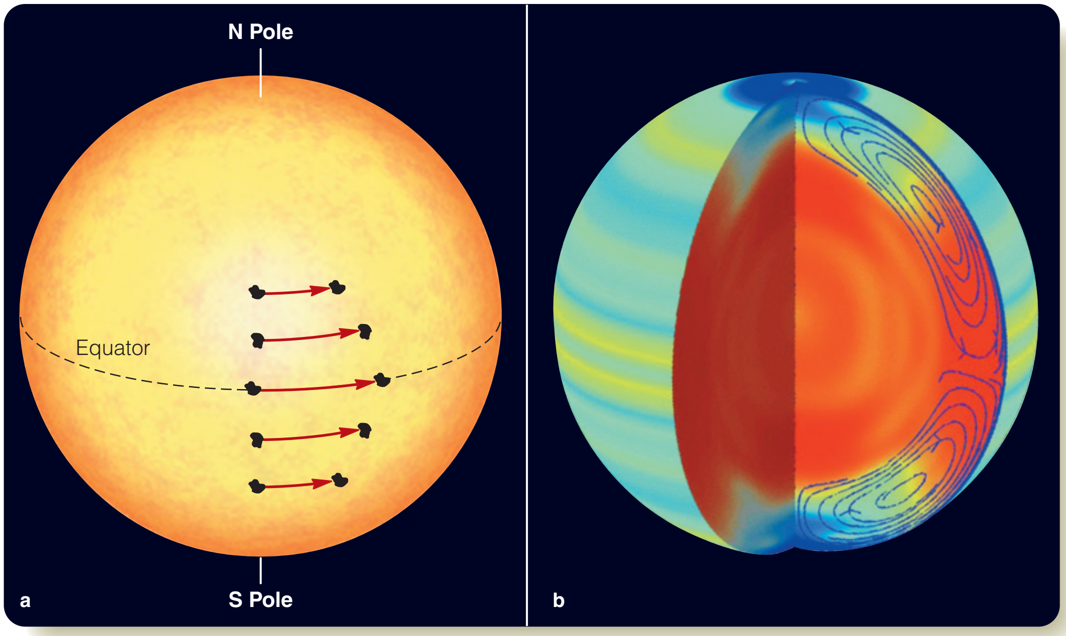
Sunspots do tend to occur in groups or pairs, and the magnetic field around the pair resembles that around a bar magnet, with one end being magnetic north and the other end magnetic south. That is just what you would expect if a magnetic tube emerged from the sun's surface through one sunspot in a pair and reentered through the other. At any one time, sunspot pairs south of the sun's equator have reversed **polarity** (orientation of their magnetic poles) compared with those north of the sun's equator. ■ Figure 7-10 illustrates this by showing sunspot pairs south of the sun's equator moving with magnetic south poles leading, and sunspots north of the sun's equator moving with magnetic north poles leading. At the end of an 11-year sunspot cycle, the spots in the new cycle appear with reversed magnetic polarities relative to the spots in the old cycle.

The Babcock model accounts for the reversal of the sun's magnetic field from cycle to cycle. As the magnetic field becomes more and more tangled, adjacent regions of the sun are dominated by magnetic fields that point in different directions. After about 11 years of tangling, the field becomes very complicated. Regions of weak north or south polarity "flip" into alignment with neighboring regions of stronger polarity. The entire field then quickly rearranges itself into a simpler pattern, the number of sunspots drops nearly to zero, and the cycle ends. Then, differential rotation begins winding up the magnetic field to start a new cycle. But the newly organized field is reversed, and the next sunspot cycle begins with the magnetic-north end of sunspot groups replaced by magnetic-south. Thus, although the cycle of numbers of sunspots is 11 years long, the complete magnetic cycle is actually 22 years long.

This magnetic cycle also explains the Maunder butterfly diagram. The Babcock model predicts that, as a sunspot cycle begins, the twisted tubes of magnetic force first begin to produce sunspot pairs at high latitudes on the sun. In other words, the first sunspots

■ **Figure 7-8**

(a) In general, the photosphere of the sun rotates faster at the equator than at higher latitudes. If you started five sunspots in a row, they would not stay lined up as the sun rotates. (b) Detailed analysis of the sun's rotation from helioseismology reveals regions of slow rotation (blue) and rapid rotation (red). Such studies show that the interior of the sun rotates differentially and that currents similar to the trade winds in Earth's atmosphere flow through the sun.



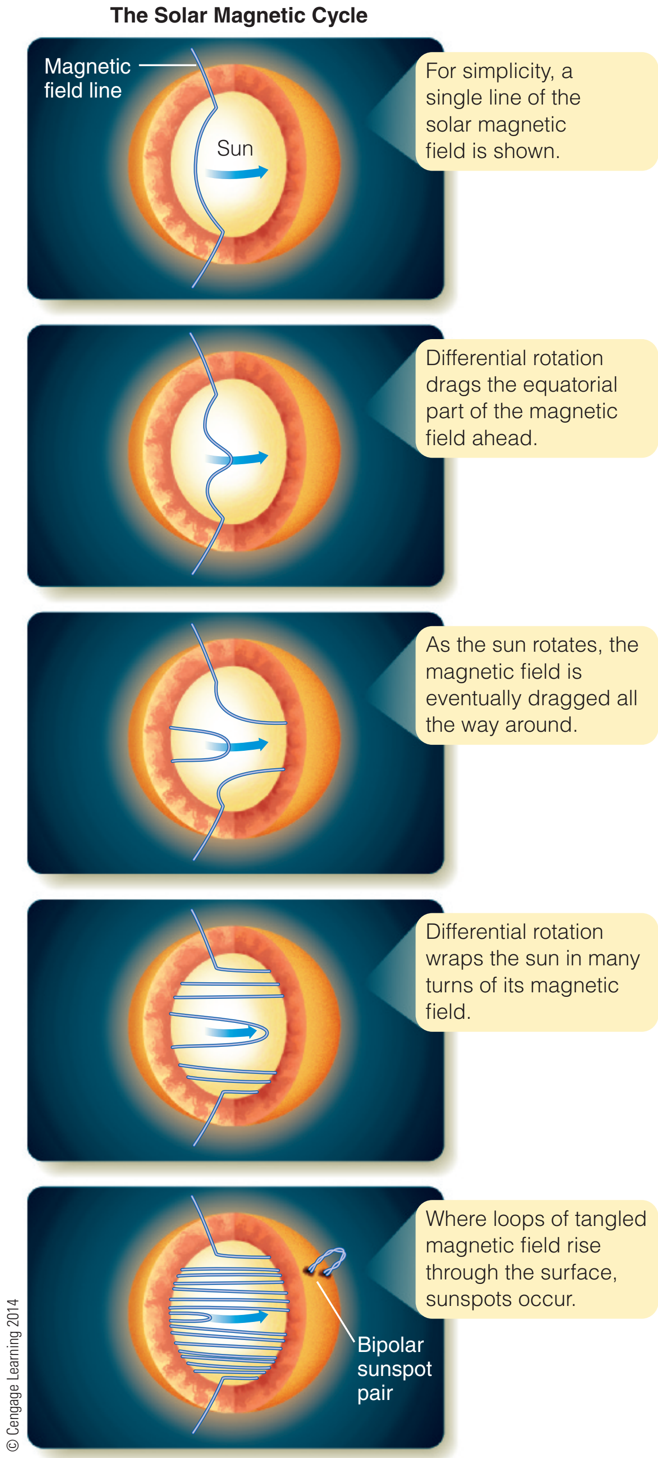


Figure 7-9

The Babcock model of the solar magnetic cycle explains the sunspot cycle as a consequence of the sun's differential rotation gradually winding up the magnetic field near the base of the sun's outer, convective layer.

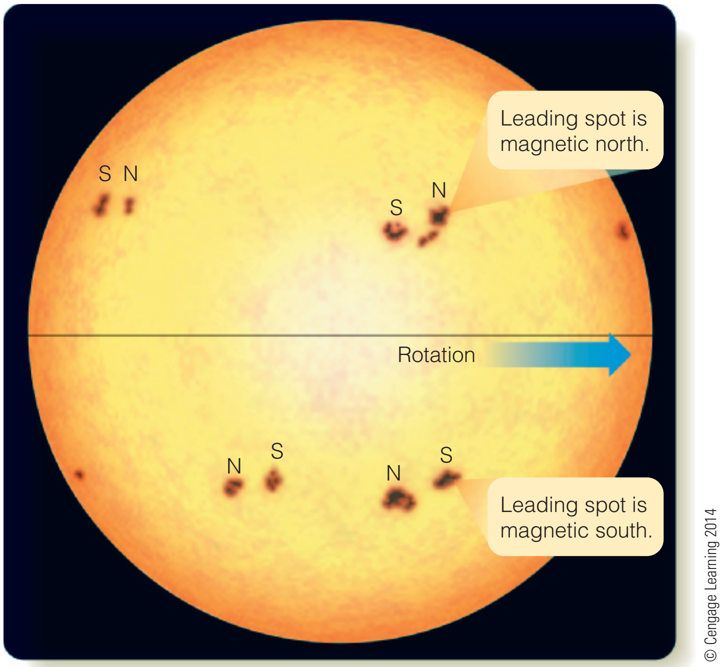


Figure 7-10

In sunspot groups, here simplified into pairs of major spots, the leading spot and the trailing spot have opposite magnetic polarity. Spot pairs in the southern hemisphere have reversed polarity from those in the northern hemisphere.

in a new cycle appear far from the sun's equator. Later in the cycle, when the field is more tightly wound, the tubes of magnetic force arch up through the surface at lower latitudes. As a result, sunspot pairs later in a cycle appear closer to the equator.

Notice the power of a scientific model. The Babcock model may be incorrect in some details, but it provides a framework around which to organize descriptions of complex solar activity. Even though the model of the sky in Chapter 2 and the model of the atom in Chapter 6 are only partially correct, they served as organizing themes to guide your thinking. Similarly, although the precise details of the solar magnetic cycle are not yet understood, the Babcock model gives you a general picture of the behavior of the sun's magnetic field (**How Do We Know? 7-1**).

Chromospheric and Coronal Activity

The solar magnetic fields extend high into the chromosphere and corona, where they produce beautiful and powerful phenomena. Study **Magnetic Solar Phenomena and the Sun–Earth Connection** on pages 132–133 and notice three important points and seven new terms:

1. All solar activity is magnetic. The arched shapes of *prominences* are produced by magnetic fields, and *filaments* are prominences seen from above.

How Do We Know? 7-1

Confirmation and Consolidation

What do scientists do all day? The scientific method is sometimes portrayed as a kind of assembly line where scientists crank out new hypotheses and then test them through observation. In reality, scientists don't often generate entirely new hypotheses. It is rare that an astronomer makes an observation that disproves a long-held theory and triggers a revolution in science. Then what is the daily grind of science really about?

Many observations and experiments confirm already-tested hypotheses. The biologist knows that all worker bees in a hive are sisters because they are all female, and they all had the same mother, the queen bee. A biologist can study the DNA from many workers and confirm that hypothesis. By repeatedly confirming a hypothesis, scientists build confidence in the hypothesis and may be able to extend it. Do all of the workers in a hive have the same father, or did the queen mate with more than one male drone?

Another aspect of routine science is consolidation, the linking of a hypothesis to other well-studied phenomena. A biologist can study yellow jacket wasps from a single nest and discover that the wasps, too, are sisters. There must be a queen wasp who lays all of the eggs in a nest. But in a few nests, the scientist may find two sets of unrelated sister workers. Those nests must contain two queens sharing the nest for convenience and protection. From his study of wasps, the biologist consolidates what he knows about bees with what others have learned about wasps and reveals something new: that bees and wasps have evolved in similar ways for similar reasons.

Confirmation and consolidation allow scientists to build confidence in their understanding and extend it to explain more about nature.


Michael Dunham/Minden Pictures/Getty Images

A yellow jacket is a wasp from a nest containing a queen wasp.

2 Tremendous energy can be stored in arches of magnetic fields, and when two arches encounter each other a *reconnection event* can release powerful eruptions called *flares.* Although these eruptions occur far from Earth, they can affect us in dramatic ways, and *coronal mass ejections (CMEs)* can trigger communications blackouts and *auroras.*

3 In some regions of the solar surface, the magnetic field does not loop back. High-energy gas from these *coronal holes* flows outward and produces much of the solar wind.

Auroras are sometimes called the "Northern Lights," but they can be viewed from high latitudes in both the Northern and Southern Hemispheres. Now, if you ever have an opportunity to watch a beautiful aurora display, you will know that you are actually seeing emission lines from gases in Earth's upper atmosphere excited to glow by a complicated interaction with the solar wind and Earth's magnetic field.

The Solar Constant

Even a small change in the sun's energy output could produce dramatic changes in Earth's climate, but humanity knows very little about the variation of the sun's energy output. The energy production of the sun can be monitored by adding up all of the energy falling on 1 square meter of Earth's surface during 1 second. Of course, you need to correct for absorption of Earth's atmosphere and also include all wavelengths from X-rays to radio waves. The result, which is called the **solar constant,** amounts to about 1360 joules per square meter per second. Is the sun really constant?

Measurements by the *Solar Maximum Mission* satellite showed variations in the energy received from the sun of about 0.1 percent that lasted for days or weeks. Superimposed on that random variation is a long-term decrease of about 0.018 percent per year that has been confirmed by observations made with other instruments. This long-term decrease may be related to a cycle of activity on the sun with a period longer than the 22-year magnetic cycle. So, careful measurements show that the solar constant is not really constant.

As you saw on page 127, the "Little Ice Age" was a period of unusually cool weather in Europe and America that lasted from about 1500 to 1850. The average temperature worldwide was about 1°C cooler than it is now. This period of cool weather included the Maunder minimum, a period of reduced solar activity—few sunspots, no auroral displays, and no solar coronas visible during solar eclipses. Scientists do not yet completely understand how those changes in the sun's surface activity would

connect to changes in Earth's average temperature. The current measured changes in the solar "constant" seemingly would cause Earth to become slightly cooler, yet our planet is clearly observed to be warming. In Chapter 9 you will learn more about the complex interaction between solar input, human activity and changes in Earth's climate.

SCIENTIFIC ARGUMENT

What kind of activity would the sun have if it didn't rotate differentially?

Once again, it can help you understand a concept by constructing an argument with one factor changed. Begin by thinking about the Babcock model. If the sun didn't rotate differentially, with its equator traveling faster than its higher latitudes, then the magnetic field would not get so tangled. As a result, there might not be a solar cycle because twisted tubes of magnetic field might not form and rise through the photosphere to produce sunspots, prominences, and flares. On the other hand, convection might still tangle the magnetic field and produce some activity. Is the magnetic activity that causes sunspots and heats the chromosphere and corona driven mostly by differential rotation, or by convection? Astronomers are not sure, but it seems likely that without differential rotation, the sun probably would not have a strong magnetic field and resulting high-temperature gas above its photosphere.

This is very speculative, but speculating within a scientific argument can be revealing. For example, redo the argument above. **How you think the sun's appearance would differ if it had no convection inside?**

7-3 Nuclear Fusion in the Sun

Like soap bubbles, stars are structures balanced between opposing forces that, if unbalanced, can destroy them. The sun is a ball of hot gas held together by its own gravity. If it were not for the sun's gravity, the hot, high-pressure gas in the sun's interior would explode outward. Likewise, if the sun were not so hot, its gravity would compress it into a small, dense body.

In this section, you will discover that the sun is powered by nuclear reactions occurring near its center. The energy released by those reactions keeps the interior hot and the gas totally ionized. That is, the temperatures are so high that none of the electrons are attached to atomic nuclei, so the gas is a soup of rapidly moving nuclei and loose electrons colliding with each other at high velocities.

How exactly can the nucleus of an atom yield energy? The answer lies in the force that holds the nuclei together.

Nuclear Binding Energy

The sun generates its energy by breaking and reconnecting the bonds between the particles *inside* atomic nuclei. There are only four different ways in which matter affects other matter, the four forces of nature: gravity, the electromagnetic force, the **weak force,** and the **strong force.** The strong force binds together atomic nuclei, and the weak force is involved in the radioactive decay of certain kinds of nuclear particles. The strong force and the weak are short-range forces that are effective only within the nuclei of atoms. Nuclear energy originates from the strong force, as nuclear reactions break and reform the bonds that hold atomic nuclei together.*

There are two ways to generate energy from atomic nuclei. Nuclear power plants on Earth use **nuclear fission** reactions that split uranium nuclei into less massive fragments. A uranium nucleus contains a total of 235 protons and neutrons, and when it decays, it splits into a range of possible fragments, each containing roughly half as many particles. Because the fragments that are produced are more tightly bound than the uranium nuclei, binding energy is released during uranium fission.

Stars don't use nuclear fission. They make energy in **nuclear fusion** reactions that combine light nuclei into heavier nuclei. The most common reaction inside stars, the one that occurs in the sun, fuses hydrogen nuclei (single protons) into helium nuclei, which contain two protons and two neutrons. Because the nuclei produced are more tightly bound than the original nuclei, net energy is released.

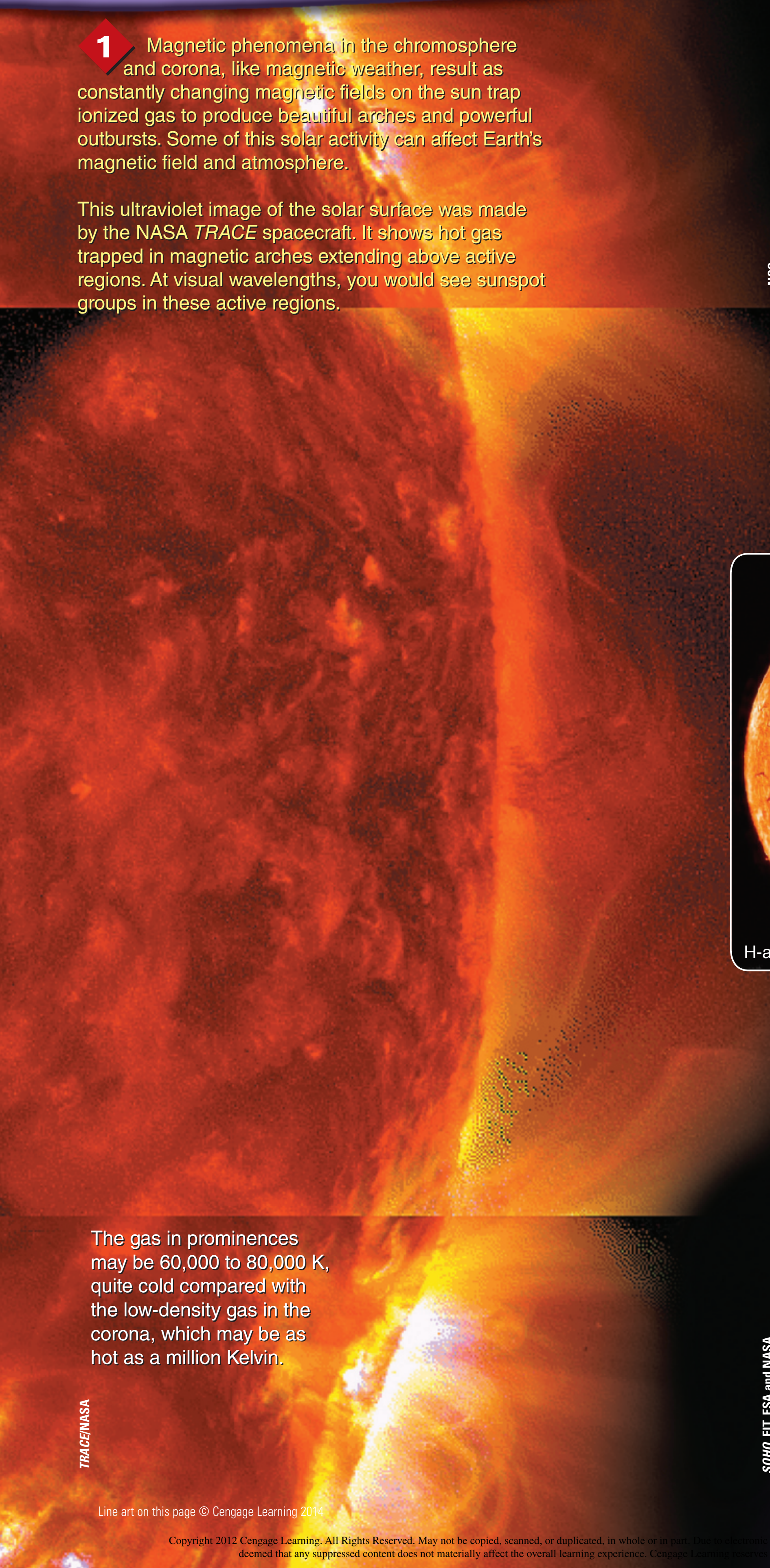
The curve plotted in ■ Figure 7-11 shows the amounts of energy that hold different atomic nuclei together. The lower in the diagram the data point for a given type of nucleus is, the more tightly the particles in that nucleus are held together. Notice that both fusion and fission reactions move downward in the diagram toward more tightly bound nuclei. Both types of reaction produce energy by releasing binding energy of atomic nuclei.

Hydrogen Fusion

The sun fuses together four hydrogen nuclei to make one helium nucleus. Because one helium nucleus has 0.7 percent less mass

*Astronomers sometimes use words like *burn* or *ignite* when they refer to processes in the core of the sun. But the nuclear reactions that power the sun and other stars are not related to ordinary burning, which is a chemical reaction. For example, the process of burning wood extracts energy by breaking and rearranging chemical bonds among atoms in the wood. Chemical bonds are formed by the electrons in atoms, and you saw in Chapter 6 that the electrons are bound to the atoms by the electromagnetic force. Therefore, the chemical energy released when those bonds are broken and rearranged comes from the electromagnetic force.

Magnetic Solar Phenomena and the Sun–Earth Connection

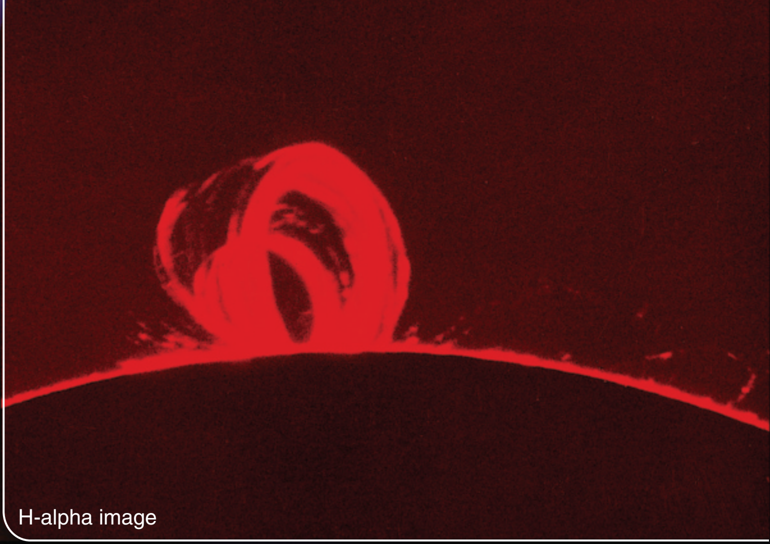


1 Magnetic phenomena in the chromosphere and corona, like magnetic weather, result as constantly changing magnetic fields on the sun trap ionized gas to produce beautiful arches and powerful outbursts. Some of this solar activity can affect Earth's magnetic field and atmosphere.

This ultraviolet image of the solar surface was made by the NASA *TRACE* spacecraft. It shows hot gas trapped in magnetic arches extending above active regions. At visual wavelengths, you would see sunspot groups in these active regions.

The gas in prominences may be 60,000 to 80,000 K, quite cold compared with the low-density gas in the corona, which may be as hot as a million Kelvin.

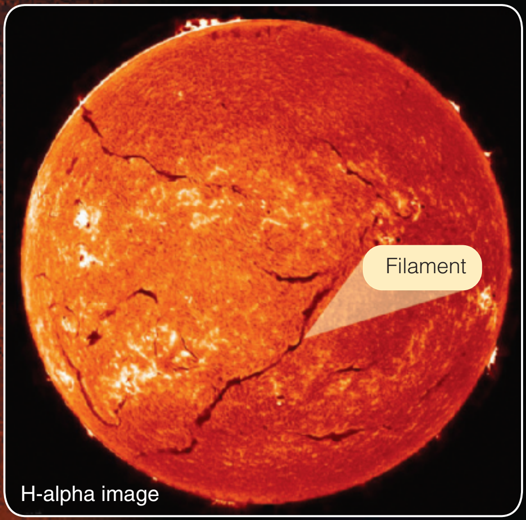
TRACE/NASA



H-alpha image

NSO

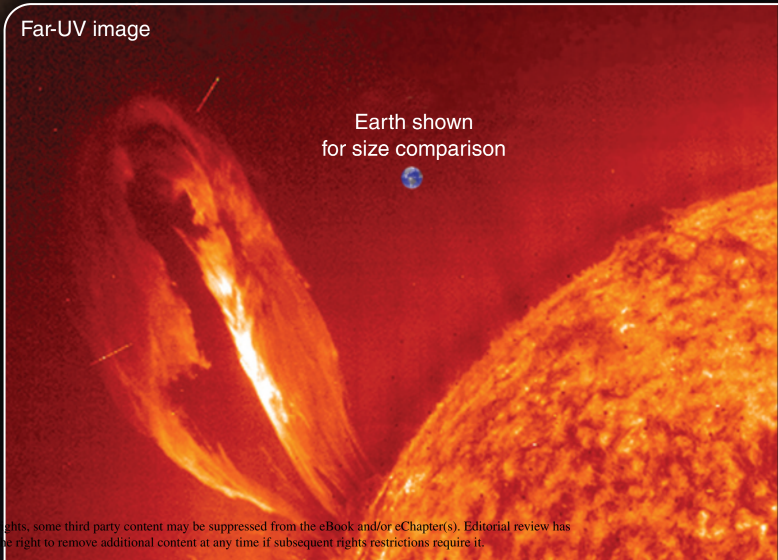
1a A **prominence** is composed of ionized gas trapped in a magnetic arch rising up through the photosphere and chromosphere into the lower corona. Seen during total solar eclipses at the edge of the solar disk, prominences look pink because of emission in the three hydrogen Balmer lines. The image above shows the arch shape suggestive of magnetic fields. Seen from above against the sun's bright surface, prominences form dark **filaments.**



H-alpha image

NOAA/SEL/USAF

1b Quiescent prominences may hang in the lower corona for many days, whereas eruptive prominences burst upward in hours. The eruptive prominence below is many Earth diameters long.

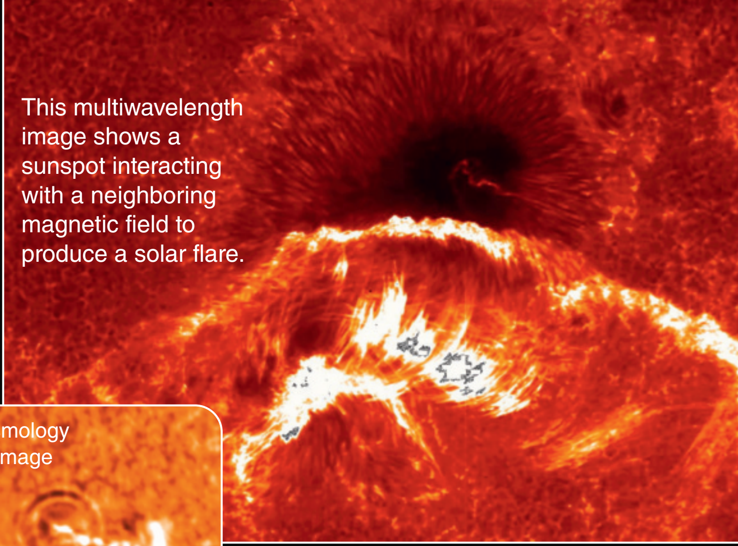


Far-UV image

SOHO, EIT, ESA and NASA

Line art on this page © Cengage Learning 2014

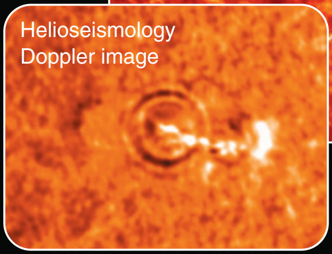
2 Solar **flares** rise to maximum in minutes and decay in an hour. They occur in active regions where oppositely directed magnetic fields meet and cancel each other out in what astronomers call **reconnection events.** Energy stored in the magnetic fields is released as short-wavelength photons and as high-energy protons and electrons. X-ray and ultraviolet photons reach Earth in 8 minutes and increase ionization in our atmosphere, which can interfere with radio communications. Particles from flares reach Earth hours or days later as gusts in the solar wind, which can distort Earth's magnetic field and disrupt navigation systems. Solar flares can also cause surges in electrical power lines and damage to Earth satellites.



This multiwavelength image shows a sunspot interacting with a neighboring magnetic field to produce a solar flare.

Hinode JAXA/NASA

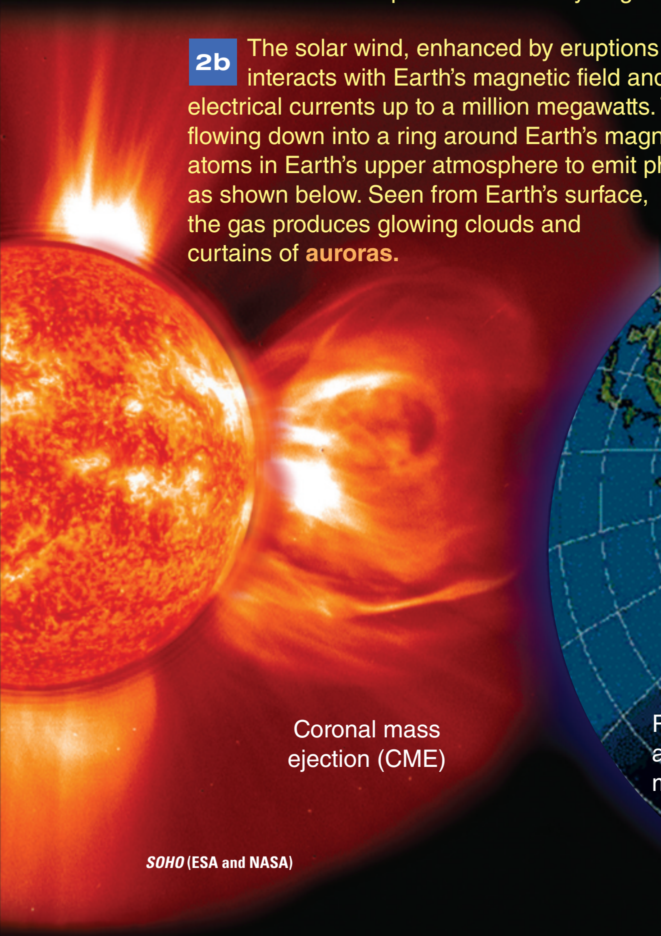
2a At right, waves rush outward at 50 km/sec from the site of a solar flare 40,000 times stronger than the 1906 San Francisco earthquake. The biggest solar flares can be a billion times more powerful than a hydrogen bomb.



Helioseismology Doppler image

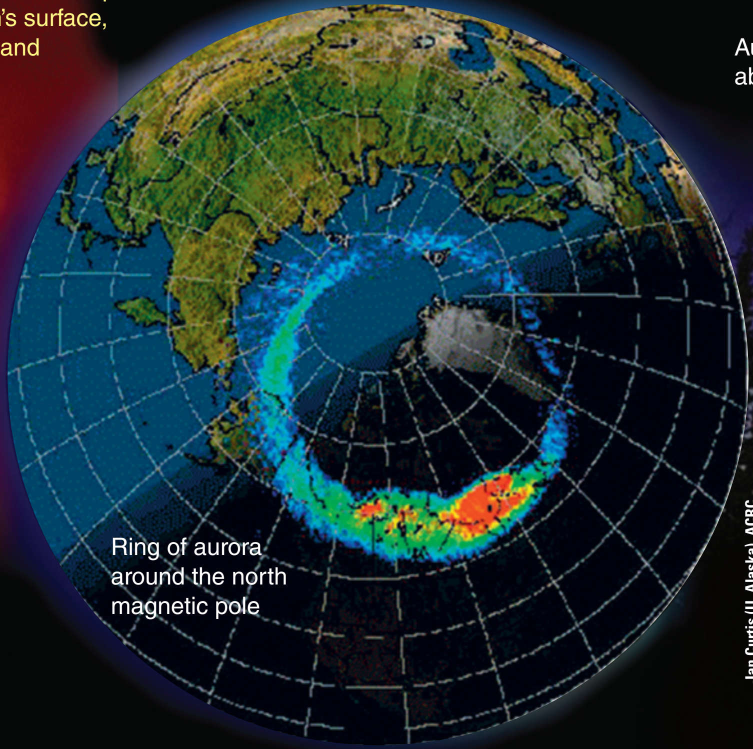
SOHO/MDI, ESA and NASA

2b The solar wind, enhanced by eruptions on the sun, interacts with Earth's magnetic field and can create electrical currents up to a million megawatts. Those currents flowing down into a ring around Earth's magnetic poles excite atoms in Earth's upper atmosphere to emit photons as shown below. Seen from Earth's surface, the gas produces glowing clouds and curtains of **auroras.**



Coronal mass ejection (CME)

SOHO (ESA and NASA)

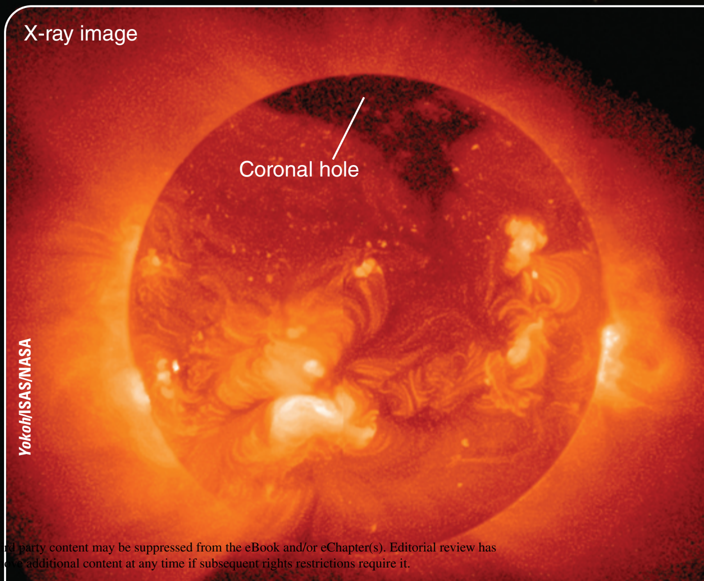


Ring of aurora around the north magnetic pole

Auroras occur about 130 km (80 mi) above the Earth's surface.

Jan Curtis (U. Alaska), ACRC

2c Magnetic reconnection events can release enough energy to blow large amounts of ionized gas outward from the corona in **coronal mass ejections (CMEs).** If a CME strikes Earth, it can produce especially violent disturbances in Earth's magnetic field.

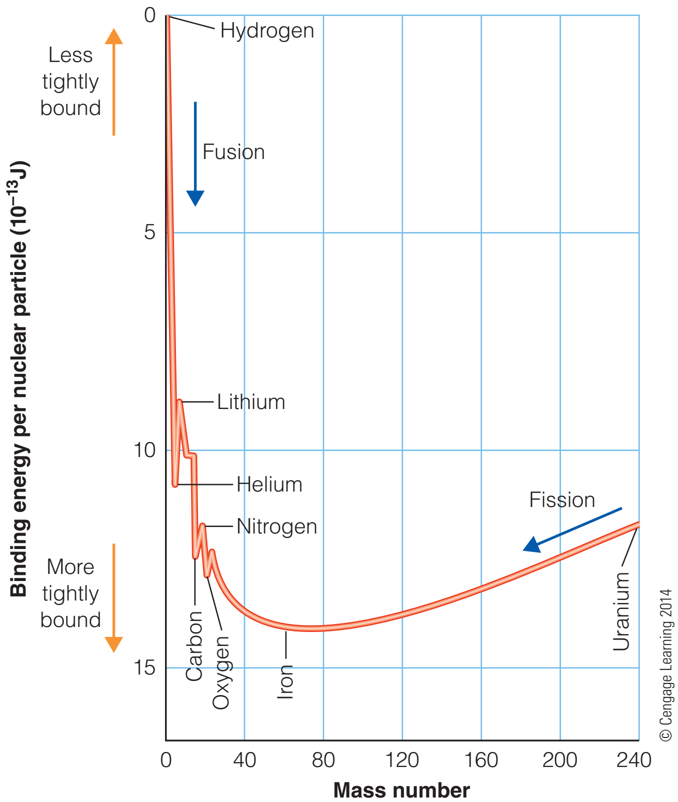


X-ray image

Yohkoh/ISAS/NASA

3 Much of the solar wind comes from **coronal holes,** where the magnetic field does not loop back into the sun. These open magnetic fields allow ionized gas in the corona to flow away as the solar wind. The dark area in this X-ray image at right is a coronal hole.

Line art on this page © Cengage Learning 2014



■ **Figure 7-11**

The orange curve in this graph shows the binding energy per particle, the energy that holds particles inside an atomic nucleus. The horizontal axis shows the atomic mass number of each element, the number of protons and neutrons in the nucleus. Both fission and fusion nuclear reactions move downward in the diagram (arrows), meaning the nucleus produced by a reaction is more tightly bound than the nuclei that went into the reaction. Iron has the most tightly bound nucleus, so no nuclear reactions can begin with iron and release energy.

Reasoning with Numbers ❙ 7-1

Hydrogen Fusion

When four hydrogen nuclei fuse to make one helium nuclei, a small amount of matter seems to disappear. To see this, subtract the mass of a helium nucleus from the mass of four hydrogen nuclei:

$$\begin{array}{r} 4 \text{ hydrogen nuclei} = 6.690 \times 10^{-27}\text{ kg} \\ -1 \text{ helium nucleus} = 6.646 \times 10^{-27}\text{ kg} \\ \hline \text{Difference in mass} = 0.044 \times 10^{-27}\text{ kg} \end{array}$$

That mass difference, 0.044×10^{-27} kg, does not actually disappear but is converted to energy according to Einstein's famous equation:

$$\begin{aligned} E &= mc^2 \\ &= (0.044 \times 10^{-27}\text{ kg}) \times (3.0 \times 10^8\text{ m/s})^2 \\ &= 4.0 \times 10^{-12}\text{ J} \end{aligned}$$

Recall that one joule is approximately equal to the energy of motion of an apple falling from a table to the floor.

than four hydrogen nuclei, it seems that some mass vanishes in the process. In fact, that mass is converted into energy, and you can figure out how much using Einstein's famous equation $E = mc^2$ (**Reasoning with Numbers 7-1**).

You can symbolize the fusion reactions in the sun with a simple equation:

$$4\,{}^1\text{H} \rightarrow {}^4\text{He} + \text{energy}$$

In that equation, ^{1}H represents a proton, the nucleus of a hydrogen atom, and ^{4}He represents the nucleus of a helium atom. The superscripts indicate the approximate mass of the nuclei, the number of protons plus the number of neutrons.

The actual steps in the process are more complicated than this convenient summary suggests. Instead of waiting for four hydrogen nuclei to collide simultaneously, a highly unlikely event, the process normally proceeds step-by-step in a series of reactions called the **proton-proton chain.** The proton-proton chain consists of three nuclear reactions that build a helium nucleus by adding together protons. Those three reactions are:

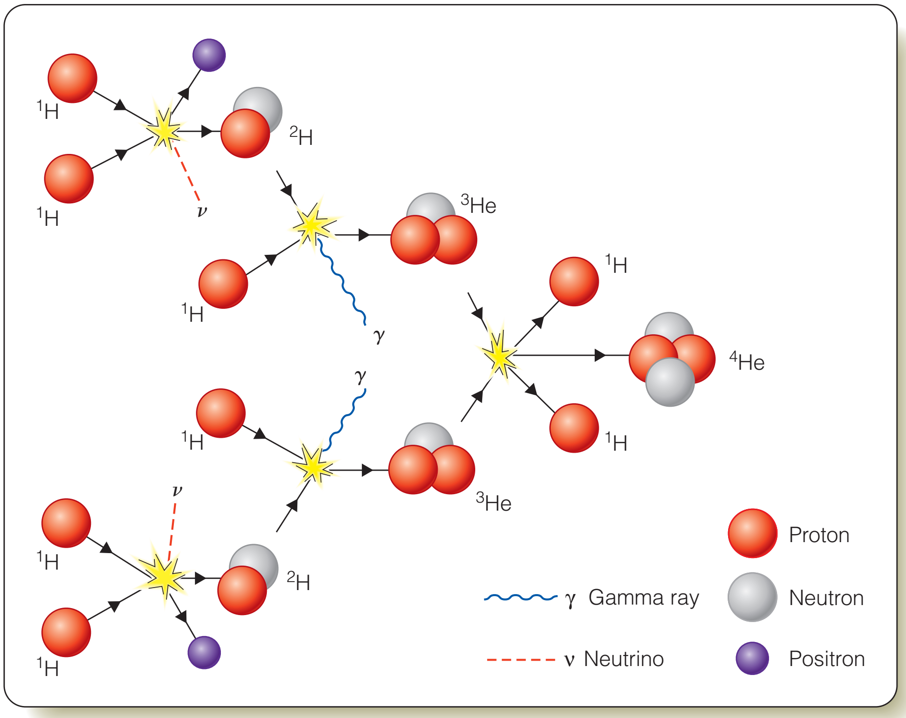
$${}^1\text{H} + {}^1\text{H} \rightarrow {}^2\text{H} + e^+ + \nu$$

$${}^2\text{H} + {}^1\text{H} \rightarrow {}^3\text{He} + \gamma$$

$${}^3\text{He} + {}^3\text{He} \rightarrow {}^4\text{He} + {}^1\text{H} + {}^1\text{H}$$

In the first reaction, two protons (two hydrogen nuclei) combine. The strong nuclear force makes the protons come together, while the weak nuclear force causes one of them to transform into a neutron and emit two particles, a **positron,** e^+ (a positively charged version of an electron), and a **neutrino,** ν (a subatomic particle having an extremely low mass and a speed nearly equal to the speed of light). Combining a proton with a neutron forms a heavy hydrogen nucleus called **deuterium.** In the second reaction, a heavy hydrogen nucleus absorbs another proton and, with the emission of a gamma-ray photon (symbolized by γ) becomes a lightweight helium nucleus. Finally, two lightweight helium nuclei combine to form a nucleus of normal helium plus two hydrogen nuclei. Because the last reaction needs two ^{3}He nuclei, the first and second reactions must occur twice (■ Figure 7-12). The net result of this chain reaction is the transformation of four hydrogen nuclei into one helium nucleus plus energy.

The energy released in the proton-proton chain appears in the form of gamma rays, positrons, neutrinos, and the energy of motion of all the particles. The gamma rays are photons that are absorbed by the surrounding gas before they can travel more than a fraction of a millimeter. That heats the gas. The positrons produced in the first reaction combine with free



© Cengage Learning 2014

■ **Figure 7-12**

The proton–proton chain combines four protons (at far left) to produce one helium nucleus (at right). Energy is produced mostly as gamma rays and as positrons, which combine with electrons and convert their mass into more energy in the form of gamma rays. Neutrinos escape without heating the gas.

electrons, and both particles vanish, converting their mass into gamma rays, which are also absorbed and help keep the gas hot. In addition, when fusion produces new nuclei, they fly apart at high velocity and collide with other particles. This energy of motion helps raise the temperature of the gas. The neutrinos, on the other hand, don't heat the gas. Neutrinos are particles that almost never interact with other particles. The average neutrino could pass unhindered through a lead wall more than a light-year thick. Consequently, the neutrinos do not warm the gas but race out of the sun at nearly the speed of light, carrying away roughly 2 percent of the energy produced by the fusion reactions.

Creating one helium nucleus makes only a small amount of energy, hardly enough to raise a housefly one-thousandth of an inch. Because one reaction produces such a small amount of energy, it is obvious that many reactions are necessary to supply the energy output of a star. The sun, for example, completes 10^{38} reactions per second, transforming 4 million tons of matter into energy every second. It might sound as if the sun is losing mass at a furious rate, but in its entire 10-billion-year lifetime, the sun will convert less than 0.07 percent of its mass into energy.

It is a **Common Misconception** that nuclear fusion in the sun is tremendously powerful. After all, the fusion of a milligram of hydrogen (roughly the mass of a match head) produces as much energy as burning 5 gallons of gasoline. However, at any one time, only a tiny fraction of the hydrogen atoms are fusing into helium, and the nuclear reactions in the sun are spread through a large volume in its core. Any single gram of matter produces only a little energy. A person of normal mass eating a normal diet produces about 3000 times more heat per gram than the matter in the core of the sun. Gram for gram, you are a much more efficient heat producer than the sun. The sun produces a lot of energy because it contains many grams of matter in its core.

Fusion reactions can occur only when the nuclei of two atoms get very close to each other. Because atomic nuclei carry positive charges, they repel each other with an electrostatic force called the Coulomb force. Physicists commonly refer to this nuclear resistance to collisions as the **Coulomb barrier.** To overcome this barrier and get close together, atomic nuclei must collide violently. Sufficiently violent collisions are rare unless the gas is very hot, in which case the nuclei move at high enough speeds. (Remember, an object's temperature is related to the speed with which its particles move.) Even so, the fusion of two protons is a highly unlikely process. If you could follow a single proton in the sun's core, you would see it encountering and bouncing off other protons millions of times a second, but you would have to follow it around for many billions of years before it happened to penetrate the Coulomb barrier and combine with another proton.

Because of the dependence of nuclear reactions on particle collisions, the reactions in the sun take place only near its center, where the gas is hot and dense. A high temperature ensures that collisions between nuclei are violent, and a high density ensures that there are enough collisions, and thus enough reactions per second, to make energy at the sun's rate. The proton–proton chain requires temperatures above about 4 million K and is vigorous only above about 10 million K.

Energy Transport in the Sun

Now you are ready to follow the energy from the core of the sun to the surface. You will learn in a later chapter that astronomers have computed models indicating that the temperature at the center of the sun must be about 16 million K for the sun to be stable. Compared with that, the sun's surface

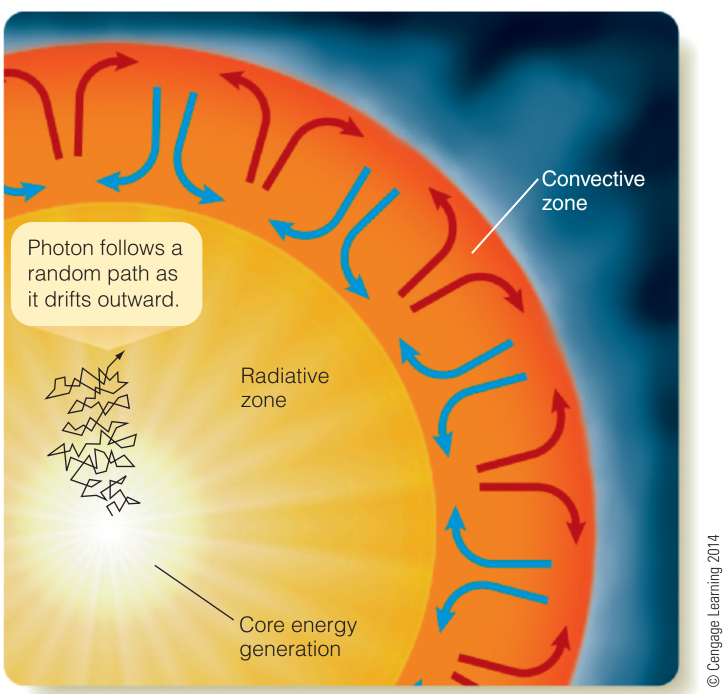
is very cool, only about 5800 K. Heat always moves from hot regions to cool regions, so energy must flow from the sun's high temperature core outward to the cooler surface where it is radiated into space.

Because the core is so hot, the photons there are gamma rays. Each time a gamma ray encounters one of the matter particles—electrons or nuclei—it is deflected or scattered in a random direction, and, as it bounces around, it slowly drifts outward toward the surface while being converted into several photons of lower energy. The net outward motion of energy is in the form of radiation, so astronomers refer to the inner parts of the sun as the **radiative zone.**

Energy originally produced in the core of the sun and traveling outward as radiation eventually reaches the outer layers of the sun where the gas is cool enough that it is not completely ionized. Such material is not very transparent to radiation. So, at that point, the energy backs up like water behind a dam, and the gas begins to churn in convection. Hot blobs of gas rise, and cool blobs sink. In this region, known as the **convective zone,** the energy is carried outward not as photons but as circulating gas (■ Figure 7-13). Rising hot gas carries energy outward, but sinking cool gas is a necessary part of the cycle. The result is net transport of energy continuing outward. Earlier in this chapter you saw that granulation and supergranulation features observed on the sun's photosphere are the visible features of energy arriving at the sun's surface from its interior.

It can take millions of years for the energy from a single gamma ray produced in the center of the sun to work its way outward first as radiation and then by convection. When the energy finally reaches the photosphere, it is radiated into space as about 2000 photons of visible light.

It is time to ask the critical question that lies at the heart of science. What is the evidence to support this theoretical explanation of how the sun shines?



■ **Figure 7-13**

A cross-section of the sun. Near the center, nuclear fusion reactions sustain high temperatures. Energy flows outward through the radiative zone as photons are randomly defected over and over by electrons. In the cooler, more opaque outer layers, the energy is carried by rising convection currents of hot gas (red arrows) and sinking currents of cooler gas (blue arrows).

Counting Solar Neutrinos

Nuclear reactions in the sun's core produce floods of neutrinos that rush out of the sun and off into space. Over 10^{14} (100 trillion) solar neutrinos flow through your body every second, but you never feel them because you are almost perfectly transparent to neutrinos. If you could detect these neutrinos, you could probe the sun's interior. You can't focus neutrinos with a lens or mirror, and they zip right through detectors used to count other atomic particles, but neutrinos of certain energies can trigger the radioactive decay of certain atoms. That gives astronomers a way to detect neutrinos.

In the 1960s, chemist Raymond Davis Jr. devised a way to count neutrinos produced by hydrogen fusion in the sun. He buried a 100,000-gallon tank of cleaning fluid (perchloroethylene, C_2Cl_4) in the bottom of a South Dakota gold mine where other types of cosmic rays could not reach it (■ Figure 7-14a) and devised a way to count individual argon atoms that were produced by neutrinos colliding with chlorine atoms in the tank.

The Davis neutrino experiment created a huge controversy. It was expected to detect one neutrino a day, but actually counted one-third as many: Only one neutrino was captured by that 100,000-gallon tank every three days. Were scientists wrong about nuclear fusion in the sun? Did they misunderstand how neutrinos behave? Was the detector not working properly? Because astronomers had reason for confidence in their understanding of the solar interior, they didn't abandon those hypotheses immediately (**How Do We Know? 7-2**). It took over 30 years, but eventually physicists

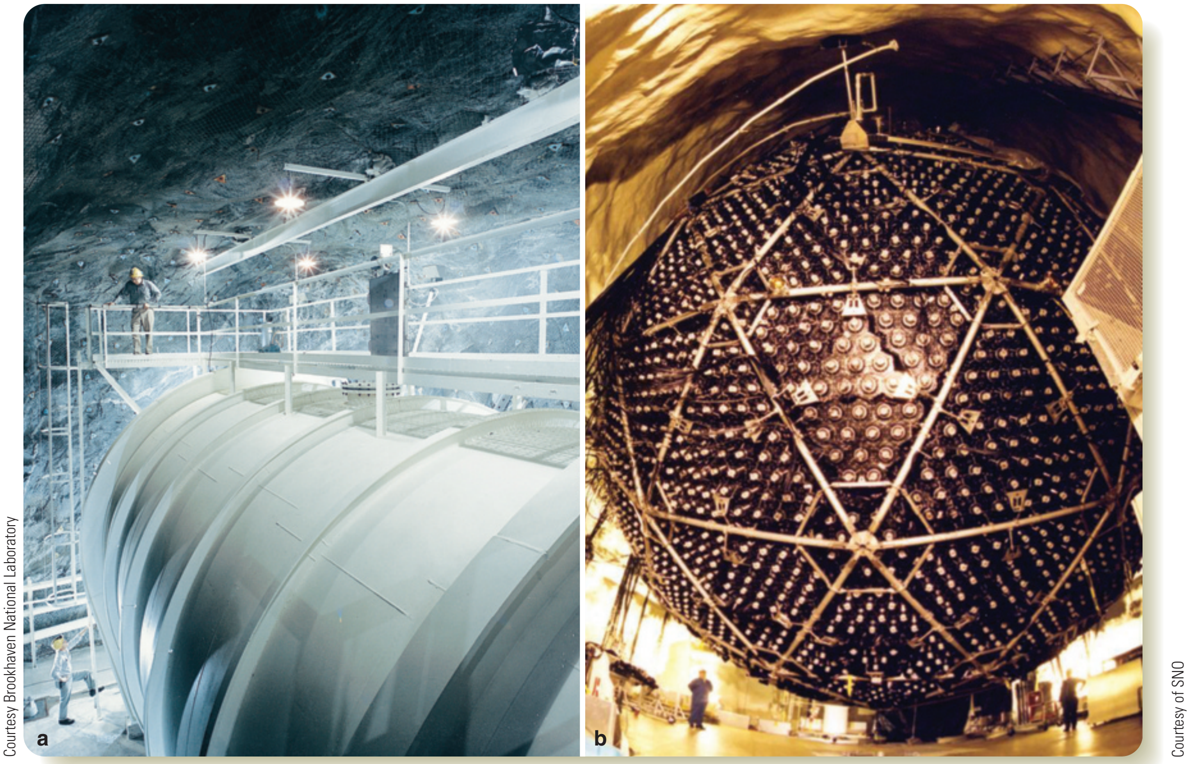


Figure 7-14

(a) The Davis solar neutrino experiment used cleaning fluid and could detect only one of the three flavors of neutrinos. (b) The Sudbury Neutrino Observatory is a 12-meter-diameter globe containing water rich in deuterium (heavy hydrogen) in place of ordinary hydrogen. Buried 2100 m (6800 ft) down in an Ontario mine, it can detect all three flavors of neutrinos and confirms that neutrinos oscillate.

were able to build better and different detectors (Figure 7-14b). They discovered that neutrinos change back and forth among three different types, which physicists call "flavors." Nuclear reactions in the sun produce only one flavor, and the Davis experiment was designed to detect (taste) that flavor. But during the 8-minute journey from the sun's core to Earth, the neutrinos changed flavor so many times that they were evenly distributed among the three different flavors by the time they arrived. That's why the Davis experiment detected only one-third of the number originally predicted. Models of nuclear fusion in the sun are now confirmed once the actual properties of neutrinos are taken into account.

The center of the sun seems forever beyond human experience, but counting solar neutrinos provides evidence to confirm the theories. The sun makes its energy through nuclear fusion.

SCIENTIFIC ARGUMENT

Why does nuclear fusion require that the gas be very hot?

This argument has to include the basic physics of atoms and thermal energy. Inside a star, the gas is so hot it is ionized, which means the electrons have been stripped off the atoms, leaving bare, positively charged nuclei. In the case of hydrogen, the nuclei are single protons. These atomic nuclei repel each other because of their positive charges, so they must collide with each other at high velocity if they are to overcome that repulsion and get close enough together to fuse. If the atoms in a gas are moving rapidly, then the gas must have high temperature, so nuclear fusion requires that the gas be very hot. If the gas is cooler than about 4 million K, hydrogen can't fuse because the protons don't collide violently enough to overcome the repulsion of their positive charges.

It is easy to see why nuclear fusion in the sun requires high temperature, but now expand your argument. **Why is fusion helped by high density?**

How Do We Know? 7-2

Scientific Confidence

How can scientists be certain of something? Sometimes scientists stick so firmly to their ideas in the face of contradictory claims that it almost sounds as if they are merely stubbornly refusing to consider alternatives. To understand what's actually going on, you might consider the perpetual motion machine, a device that supposedly runs continuously with no source of energy. If you could find a real perpetual motion machine, you could make cars that would run without any fuel. That's good mileage.

For centuries many people have each claimed to have invented perpetual motion machines, and for just as long scientists have been dismissing these claims as impossible. The problem with a perpetual motion machine is that it violates the law of conservation of energy, and scientists are not willing to accept that the law could be wrong. In fact, the Royal Academy of Sciences in Paris was so sure that a perpetual motion machine was impossible, and so tired of debunking hoaxes, that in 1775 they issued a formal statement refusing to deal with them. The U.S. Patent Office policy is that they won't even consider starting the patent process for one without seeing a working model first. Why do scientists seem so stubborn and close-minded on this issue?

Why isn't one person's belief in perpetual motion just as valid as another person's belief in the law of conservation of energy? In fact, the two positions are not equally valid. The confidence physicists have in that law is not a belief or even an opinion; it is an understanding founded on the fact that the law has been tested uncountable times and has never failed. In contrast, no one has ever successfully demonstrated a perpetual motion machine. The law of conservation of energy is a fundamental truth about nature and can be used to understand what is possible and what is impossible.

When the first observations of solar neutrinos detected fewer than were predicted, some scientists speculated that astronomers misunderstood how the sun makes its energy or that they misunderstood the internal structure of the sun. But astronomers stubbornly refused to reject their model because the nuclear physics of the proton-proton chain is well understood, and models of the sun's structure have been tested successfully many times. The confidence astronomers felt in their understanding of the sun was an example of scientific certainty, and that confidence in basic natural laws prevented them from abandoning decades of work in the face of a single contradictory observation.

What seems to be stubbornness among scientists is really their confidence in basic principles that have been tested over and over. Those principles are the keel that keeps the ship of science from rocking before every little breeze. Without even looking at that perpetual motion machine, your physicist friends can warn you not to invest.

© Cengage Learning 2014

For centuries, people have tried to design a perpetual motion machine, but not a single one has ever worked. Scientists understand why.

What Are We? Sunlight

We live very close to a star and depend on it for survival. All of our food comes from sunlight that was captured by plants on land or by plankton in the oceans. We either eat those plants directly or eat the animals that feed on those plants. Whether you had salad, seafood, or a cheeseburger for supper last night, you dined on sunlight, thanks to photosynthesis.

Almost all of the energy that powers human civilization came from the sun through photosynthesis in ancient plants that were buried and converted to coal, oil, and natural gas. New technology is making energy from plant products like corn, soy beans, and sugar. It is all stored sunlight. Windmills generate electrical power, and the wind blows because of heat from the sun. Photocells make electricity directly from sunlight. Even our bodies have adapted to use sunlight to help manufacture vitamin D.

Our planet is warmed by the sun, and without that warmth the oceans would be ice and much of the atmosphere would be a coating of frost. Books often refer to the sun as "our sun" or "our star." It is ours in the sense that we live beside it and by its light and warmth, but we can hardly say it belongs to us. It is closer to the truth to say that we belong to the sun.

Study and Review

Summary

- The sun's visible surface, the **photosphere (p. 118),** is the layer in the sun from which visible photons most easily escape. The solar atmosphere consists of two layers of hot, low-density gas above the photosphere: the **chromosphere (p. 118)** and **corona (p. 118).**
- The **granulation (p. 119)** of the photosphere is produced by **convection (p. 119)** currents of hot gas rising from below. **Supergranules (p. 119)** appear to be caused by larger convection currents deeper in the sun.
- The edge or limb of the solar disk is dimmer than the center. This limb darkening is evidence that the temperature in the solar atmosphere increases with depth.
- The chromosphere is most easily visible during total solar eclipses, when it flashes into view for a few seconds. It is a thin, hot layer of gas just above the photosphere. Its pink color is caused by the hydrogen Balmer emission lines in its spectrum.
- **Filtergrams (p. 121)** of the chromosphere reveal **spicules (p. 121),** flamelike structures extending upward into the lower corona.
- The corona is the sun's outermost atmospheric layer and can be imaged using a **coronagraph (p. 121).** It is composed of a very low-density, very hot gas extending many solar radii from the visible sun. Its high temperature—up to 2 million K—is understood to be maintained by energy transported via motions of the magnetic field extending up through the photosphere—the **magnetic carpet (p. 122)**—and by magnetic waves coming from below the photosphere.
- Parts of the corona give rise to the **solar wind (p. 122),** a breeze of low-density ionized gas streaming away from the sun.
- The strength of spectral lines in the sun's spectrum depends partly on the temperature of its atmosphere. Some elements that are abundant have weak spectral lines, and some that are not very abundant have strong spectral lines. The mathematical techniques for deriving true abundances of elements from the solar spectrum were worked out by Cecilia Payne in the 1920s.
- Solar astronomers can study the motion, density, and temperature of gases inside the sun by analyzing the way the solar surface oscillates. Known as **helioseismology (p. 123),** this field of study requires large amounts of data and extensive computer analysis.
- The sun's light and infrared radiation can burn your eyes, so you must take great care in observing it. Dark **sunspots (p. 118)** come and go on the sun, but only rarely are they large enough to be visible to the unaided eye.
- Sunspots seem dark because they are slightly cooler than the rest of the photosphere. The average sunspot is about twice the size of Earth. They appear for a month or so and then fade away, and the number of spots on the sun varies with an 11-year cycle.
- Early in a sunspot cycle, spots appear farther from the sun's equator, and later in the cycle they appear closer to the equator. This is shown in the **Maunder butterfly diagram (p. 126).**
- Astronomers can use the **Zeeman effect (p. 127)** to measure magnetic fields on the sun. The average sunspot contains magnetic fields a few thousand times stronger than Earth's field. This is part of the evidence that the sunspot cycle is produced by a solar magnetic cycle that is 22 years long.
- The sunspot cycle does not repeat exactly each cycle, and the decades from 1645 to 1715, known as the **Maunder minimum (p. 127),** seem to have been a time when solar activity was very low and Earth's climate was slightly colder.
- Sunspots are the visible consequences of **active regions (p. 127)** where the sun's magnetic field is strong. Arches of magnetic field can produce sunspots where the field passes through the photosphere.
- The sun's magnetic field is produced by the **dynamo effect (p. 128)** operating at the base of the zone of convection currents.
- Alternate sunspot cycles have reversed magnetic **polarity (p. 128),** which has been explained by the **Babcock model (p. 128),** in which the sun's **differential rotation (p. 128)** and convection currents tangle the magnetic field. The magnetic field's tangles arch through the photosphere and cause active regions visible to your eyes as sunspot pairs. When the field becomes strongly tangled, it reorders itself into a simpler but reversed field, and the cycle starts over.
- Arches of magnetic field are visible as **prominences (p. 132)** in the chromosphere and corona. Seen from above in filtergrams, prominences are visible as dark **filaments (p. 132)** silhouetted against the bright chromosphere.
- Magnetic field **reconnection events (p. 133)** can produce powerful **flares (p. 133),** sudden eruptions of X-ray, ultraviolet, and visible radiation plus high-energy atomic particles. Flares are important because they can have dramatic effects on Earth, such as communications blackouts.
- The solar wind originates in regions on the solar surface called **coronal holes (p. 133),** where the sun's magnetic field leads out into space and does not loop back to the sun. **Coronal mass ejections,** or **CMEs (p. 133),** occur when magnetic fields on the surface of the sun eject bursts of ionized gas that flow outward in the solar wind. Such bursts can produce **auroras (p. 133)** and other phenomena if they strike Earth.
- There are only four forces in nature: the electromagnetic force, the gravitational force, the **weak force (p. 131),** and the **strong force (p. 131).** The energy in nuclear reactions comes from the strong force.
- Nuclear reactors on Earth generate energy through **nuclear fission (p. 131),** in which large nuclei such as uranium break into smaller fragments. The sun generates its energy through **nuclear fusion (p. 131),** in which hydrogen nuclei fuse to produce helium nuclei.
- Hydrogen fusion in the sun proceeds in three steps known as the **proton-proton chain (p. 134).** The first step in the chain combines two hydrogen nuclei to produce a heavy hydrogen nucleus called **deuterium (p. 134).** The second step produces light helium, and the third step combines the light helium nuclei to form normal helium. Energy is released as gamma rays, in the form of **positrons (p. 134)** and **neutrinos (p. 134),** and as rapid motion of those particles and atomic nuclei.
- Fusion can occur only at the center of the sun because charged particles repel each other, and high temperatures are needed to give particles high enough velocities to penetrate this **Coulomb barrier (p. 135).** High densities are needed to provide large numbers of reactions.
- Energy flows out of the sun's core as photons traveling through the **radiative zone (p. 136)** and closer to the surface as rising currents of hot gas and sinking currents of cooler gas in the **convective zone (p. 136).**

- Neutrinos escape from the sun's core at nearly the speed of light, carrying away about 2 percent of the energy produced by fusion. Observations of fewer neutrinos than expected coming from the sun's core are now explained by the oscillation of neutrinos among three different types (flavors). The detection rate of solar neutrinos confirms details of the hypothesis that the sun's energy comes from the proton-proton chain of hydrogen fusion reactions.

Review Questions

1. Why can't you see deeper into the sun than the photosphere?
2. What evidence can you give that granulation is caused by convection?
3. How are granules and supergranules related? How do they differ?
4. How can astronomers detect structure in the chromosphere?
5. What evidence can you give that the corona has a very high temperature?
6. What heats the chromosphere and corona to a high temperature?
7. Why does hydrogen, which is abundant in the sun's atmosphere, have relatively weak spectral lines, while calcium, which is not abundant, has very strong spectral lines?
8. How are astronomers able to explore the layers of the sun below the photosphere?
9. What evidence can you give that sunspots are magnetic?
10. How does the Babcock model explain the sunspot cycle?
11. What does the spectrum of a prominence reveal? What does its shape reveal?
12. How can solar flares affect Earth?
13. Why does nuclear fusion require high temperatures?
14. Why does nuclear fusion in the sun occur only near the center?
15. How can astronomers detect neutrinos from the sun?
16. How did neutrino oscillation affect the detection of solar neutrinos by the Davis experiment?
17. **How Do We Know?** What does "confirmation and consolidation of facts" mean? How does that extend scientific understanding?
18. **How Do We Know?** What does it mean when scientists say they are certain? What does scientific certainty really mean?

Discussion Questions

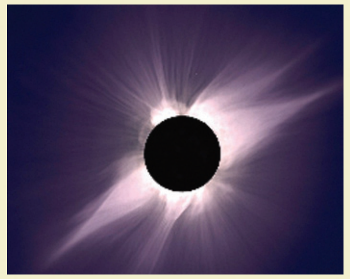
1. Explain why the presence of spectral lines of a given element in the solar spectrum tells you that element is present in the sun, but the absence of the lines would not necessarily mean the element is absent from the sun.
2. What energy sources on Earth cannot be thought of as stored sunlight?
3. What would the spectrum of an auroral display look like? Why?
4. What observations would you make if you were ordered to set up a system that could warn astronauts in orbit of dangerous solar flares? (Such a warning system actually exists.)

Problems

1. The radius of the sun is 0.7 million km. What fraction of the radius is taken up by the chromosphere?
2. The smallest detail visible with ground-based solar telescopes is about 1 arc second. How large a region does this represent on the sun? (*Hint:* Use the small-angle formula, Chapter 3.)
3. What is the angular diameter of a star like the sun located 5 ly from Earth? Is the *Hubble Space Telescope* able to detect detail on the surface of such a star? (*Hint:* Use the small-angle formula, Chapter 3.)
4. If a sunspot has a temperature of 4200 K and the average solar photosphere has a temperature of 5800 K, how many times brighter is a square meter of the photosphere compared to a square meter of the sunspot? (*Hint:* Use the Stefan-Boltzmann law, Chapter 6.)
5. How much energy is produced when the sun converts 1 kg of mass into energy?
6. How much energy is produced when the sun converts 1 kg of hydrogen into helium? (*Hint:* How does this problem differ from Problem 5?)
7. A 1-megaton nuclear weapon produces about 4×10^{15} J of energy. How much mass must vanish when a 5-megaton weapon explodes?
8. A solar flare can release 10^{25} J. How many megatons of TNT would be equivalent? (See Problem 7 for conversion between megatons and joules.)
9. The United States consumes about 2.5×10^{19} J of energy in all forms in a year. How many years could you run the United States on the energy released by the solar flare in Problem 8?
10. Use the luminosity of the sun, the total amount of energy it emits each second, to calculate how much mass it converts to energy each second.

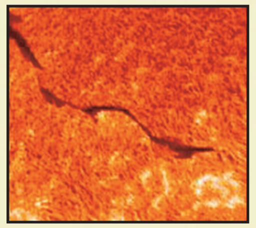
Learning to Look

1. Whenever there is a total solar eclipse, you can see something like the image shown below. Explain why the shape and extent of the glowing gases is observed to be different for each eclipse.

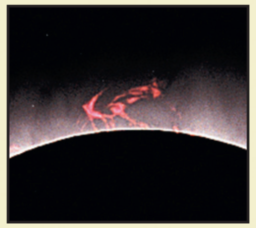


NOAO

2. The two images here show two solar phenomena. What are they, and how are they related? How do they differ?

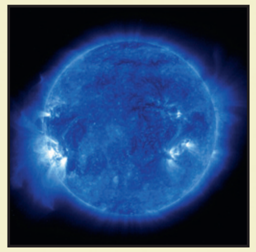


NOAO



NOAO

3. This image of the sun was recorded in the extreme ultraviolet by the *SOHO* spacecraft. Explain the features you see.



SOHO/ESA & NASA

Great Debates

1. *Do Humans Need the Sun?* Think about your daily life on Earth and the sun's effect on Earth. With today's technology, can humans survive without the sun's energy?
   a. Use at least three vocabulary words about the sun correctly in your two debates, underlining the vocabulary words.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

2. *The Sun's Color.* The average solar photosphere has a temperature of about 5800 K. Using Figure 6-6 and this temperature, what is the peak color of the photosphere's blackbody emission? When you were a child, did you color the sun this peak color? You may have entered this class thinking the sun emits a peak color other than your answer. Now that you have corrected this internal bias, should you correct others on the peak color of the sun to alter their internal bias? For example, should you correct a child's crayon choice for color of the sun if the child did not choose this color? If you are at your employer's holiday party and a colleague comments on the color of the Sun and this color is not the peak color, should you educate your colleague on the sun's true peak emission color?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

3. *GPS and the Sun.* When solar flares and coronal mass ejections occur, especially during solar maxima, radio communications may be affected. The GPS device you use to find your current location, to navigate while driving or flying, may suddenly be disrupted. If you crash your car because the mapping tool or navigation app on your cell phone or GPS device ceases to work, is the GPS provider at fault?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

4. *Space Mirrors.* A candidate states during a debate that if she is elected president of the United States, she will install space mirrors to light up the highways at night across the country. What would be the repercussions of such an act? Should any government interfere with nature? Are space mirrors a wise use of taxpayer dollars? What effect would such a project have on astronomical observatories and amateur stargazers?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

5. *Breast Cancer/Solar Minimum Connection.* A scientific paper states that female grandchildren of females who were exposed to cosmic radiation during a solar minimum have an increased risk of breast cancer (D. Juckett, 2007, *Journal of Astrobiology*, 6: 307). The paper cites analysis of 140 years of breast cancer data. Is the paper a hypothesis, theory, model, or law? If you know a female who was born in 2007, do you inform her parents that their baby's female grandchildren have an increased risk of cancer?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Zeeman Effect
- The Proton-Proton Chain
- The Sun

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 10: Helioseismology

You might think that it would be impossible to be sure about conditions inside the sun. Your view stops at the photosphere, where the sun's atmosphere becomes dense enough to be opaque. Nevertheless, as you learned in this chapter and previous chapters, observations from Earth allow you learn a lot about the sun. Without peering inside you can measure or infer the sun's diameter, mass, atmospheric composition, total energy output, photosphere temperature, rotation speed versus latitude, magnetic field strength, and the presence of convection below the photosphere. However, until recently, in considering the sun's interior, astronomers were limited to constructing mathematical models consistent with observations but difficult to test and verify.

Helioseismology is the name for studying the sun ("helio") by analyzing the characteristics of vibrations that travel in its interior ("seismology"), analogous to studying the interior of Earth by analyzing earthquake waves. The speed of such waves depends on their wavelength and the density and temperature of the material through which they pass. Seismic waves tend to change direction when they move between regions of significantly different properties, as a light wave bends when it passes from air into glass on a path not perpendicular to the boundary. And, for solar and earthly seismic waves, as well as for electromagnetic radiation, the amount of bending depends on the wavelength.

Of course, no seismograph has been placed on the sun. Solar seismic vibrations can be detected from Earth by measuring the very small Doppler shifts of regions of the photosphere rising and falling as the waves pass underneath or arrive at the surface from somewhere across the sun's bulk. By accumulating and analyzing huge amounts of data gathered over decades, solar astronomers have precisely measured the size of the convective zone and also found that differential rotation seen on the surface extends into the interior. These results have in some cases verified and refined, but in other cases contradicted, the existing solar interior models, forcing the modelers to go back to the drawing board. To older scientists, the amount of information available now about the inside of the sun seems almost miraculous.

There is one other way to investigate the sun's interior, specifically the core region where nuclear fusion reactions take place. As you learned in this chapter, neutrinos produced by fusion come directly to us from the sun's core. Counting the neutrinos lets us monitor the nuclear reaction rate. The models can now be "tweaked" to agree with both the helioseismological data and the neutrino data.

Virtual Astronomy Lab 10, "Helioseismology," helps you make the connection between what you know about the exterior of the sun and what can be determined about its interior. You can access that lab at http://cengage.com/someaddress/VAL/Lab10.

Virtual Astronomy Lab 4: Solar Wind

Space is full, not empty. In this chapter you have learned that the outer part of the sun's atmosphere, called the corona, is extremely hot, with temperatures ranging up to several million degrees Kelvin. That high temperature means the particles in the corona are moving faster than solar escape velocity and stream away from the sun, filling interplanetary space. That flow is called the solar wind. In a sense, Earth and the other planets orbiting in the solar wind are within an extension of the corona. The outer boundary of the solar system can be considered to be the region far beyond Neptune's orbit where the solar wind collides with the gas of interstellar space.

The cause of the corona's high temperature is not well understood, although, as you also learned in this chapter, heating by some mechanism that delivers magnetic energy from the convection zone below the photosphere to the corona is probably involved. At that temperature the atoms of the wind are completely or almost completely ionized. Solar observatories on the ground and in orbit have found that the solar wind does not flow uniformly from the sun in all directions, nor is it a steady breeze. Instead there are coronal "holes" where the magnetic field lines are not looped but open so that the ionized coronal gas is guided away from the sun, and CME (coronal mass ejection) events that are like "gusts" in the wind.

Sections 1 and 2 of Virtual Astronomy Lab 4, "Solar Wind and Cosmic Rays," guide you to further understanding about the corona and the solar wind's speed, mass flow rate, composition, and interaction with Earth's magnetic field. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

8

Origin of the Solar System and Extrasolar Planets

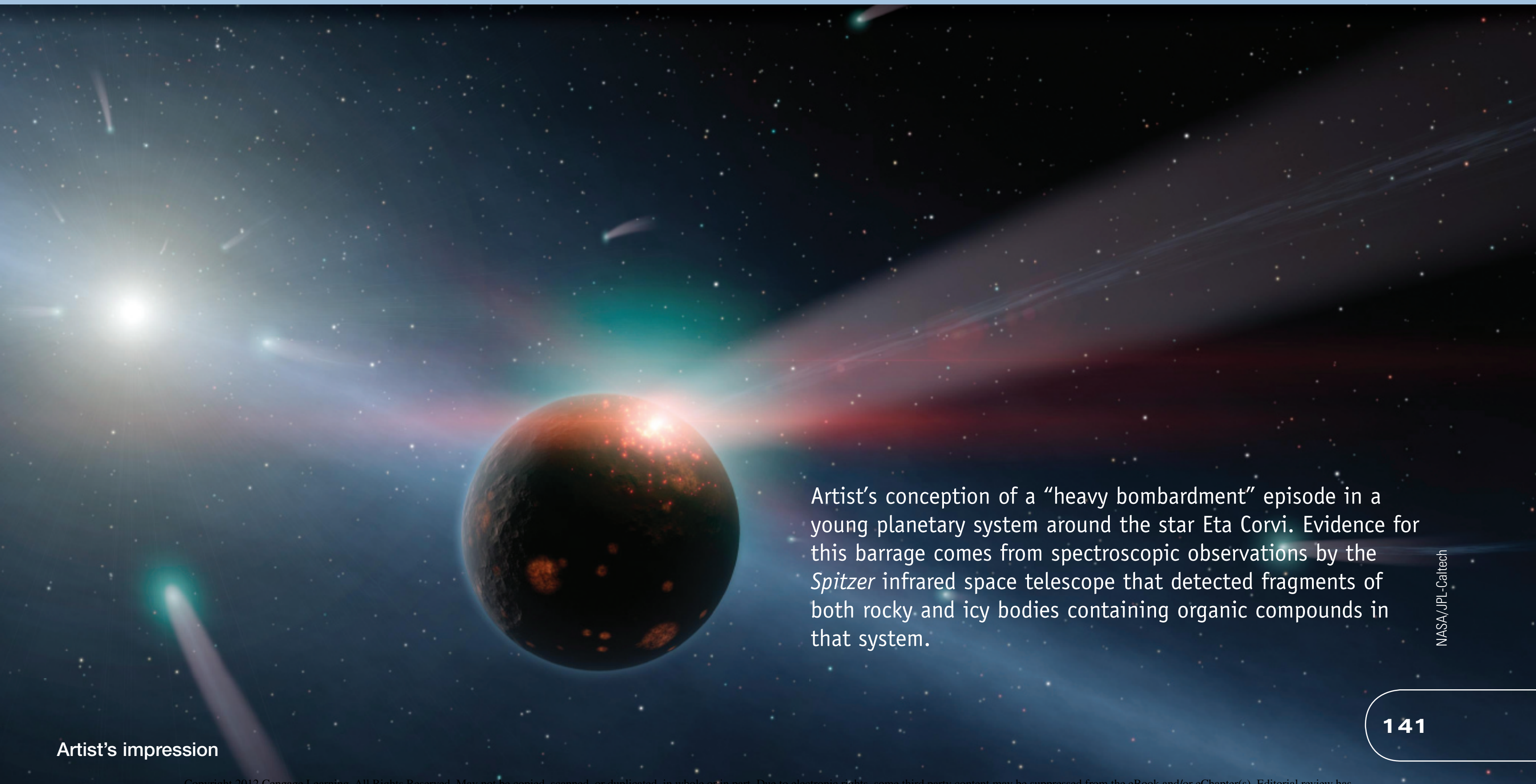
Guidepost

You have studied the appearance, origin, structure, and evolution of stars, galaxies, and the entire universe. So far, though, your studies have left out one important type of object—planets. Now it is time for you to fill in that blank. In this chapter, once you learn the evidence for how the solar system formed, you can understand how the processes you have been studying produced Earth, your home planet.

As you explore our solar system in space and time, you will find answers to four important questions:

- **What are the observed properties of the solar system as a whole?**
- **How did the solar system form?**
- **How do planets form?**
- **What do astronomers know about other planetary systems?**

In the following four chapters you will explore in more detail each of the planets, plus asteroids, comets, and meteorites. By studying the solar system as a whole, and its origin, before studying the individual objects in it, you give yourself a better framework for understanding these fascinating worlds.



Artist's conception of a "heavy bombardment" episode in a young planetary system around the star Eta Corvi. Evidence for this barrage comes from spectroscopic observations by the *Spitzer* infrared space telescope that detected fragments of both rocky and icy bodies containing organic compounds in that system.

What place is this?
Where are we now?

CARL SANDBURG,
"GRASS"

MICROSCOPIC CREATURES LIVE in the roots of your eyelashes. Don't worry. Everyone has them, and they are harmless.* They hatch, fight for survival, mate, lay eggs, and die in the tiny spaces around the roots of your eyelashes without doing any harm. Some live in renowned places—the eyelashes of a glamorous movie star, for example—but the tiny beasts are not self-aware; they never stop to say, "Where are we?"

You can study the solar system for many reasons. You can study Earth and its sibling planets because, as you are about to discover, there are almost certainly more planets in the universe than stars. Above all, you can study the solar system because it is your home in the universe. Because humans are an intelligent species, we have the ability and the responsibility to wonder where we are and what we are. Our kind have inhabited this solar system for at least a hundred thousand years, but only within the last few hundred years have we begun to understand what the solar system is.

8-1 The Great Chain of Origins

YOU ARE LINKED through a great chain of origins that leads backward through time to the instant when the universe began, 13.7 billion years ago. The gradual discovery of the links in that chain has been one of the most exciting adventures of the human intellect. In earlier chapters, you studied some of that story: the origin of the universe in the big bang, the formation of galaxies, the origin of stars, and the growth of the chemical elements. Now you will explore further to consider the origin of planets.

History of the Atoms in Your Body

The universe began in the big bang (discussed in Chapter 19). By the time the universe was a few minutes old, the protons, neutrons, and electrons in your body had come into existence. You are made of very old matter.

Although those particles formed quickly, they were not linked together to form many of the atoms that are common today. The matter in the early universe was about 75 percent hydrogen (by mass) and 25 percent helium. Although your body does not contain helium, it does contain many of those ancient hydrogen atoms unchanged since the universe began. Evidence indicates that almost no atoms heavier than helium were made in the big bang.

Within a few hundred million years after the big bang, matter began to collect to form galaxies containing billions of stars. You have learned that nuclear reactions inside stars are where low-mass atoms such as hydrogen are combined to make heavier atoms (to be discussed in Chapter 15). Generation after generation of stars cooked the original particles, fusing them into atoms such as carbon, nitrogen, and oxygen that are common in your body. Even the calcium atoms in your bones were assembled inside stars.

Massive stars produce iron in their cores, but much of that iron is destroyed when the core collapses and the star explodes as a supernova. Most of the iron on Earth and in your body was produced instead by carbon fusion in type Ia supernova explosions and by the decay of radioactive atoms in the expanding matter ejected by type II supernovae. Atoms heavier than iron such as gold, silver, and iodine are created by rapid nuclear reactions that can occur only during supernova explosions. Iodine is critical to the function of your thyroid gland, and you probably have gold and silver jewelry or dental fillings. Realize that these types of atoms, which are part of your life on Earth, were made during the violent deaths of massive stars long ago.

Our galaxy contains at least 100 billion stars, of which the sun is one. Astronomers have a variety of evidence that the sun formed from a cloud of gas and dust about 5 billion years ago, and the atoms in your body were part of that cloud. The stories of how the cloud gave birth to the planets and how the atoms in your body found their way onto Earth and into you make up the story of this chapter. As you explore the origin of the solar system, keep in mind the great chain of origins that created the atoms. As the geologist Preston Cloud remarked, "Stars have died that we might live."

The Solar Nebula Theory

Over the last two centuries, astronomers have proposed two kinds of hypotheses for the origin of the planets in our solar system. **Catastrophic hypotheses** proposed that the planets formed from some improbable event such as the collision of the sun and another star. **Evolutionary hypotheses** proposed that the planets formed gradually and naturally as the sun formed. Since around the middle of the 20th century, evidence has become overwhelming for the evolutionary scenario (**How Do We Know 8-1**). In fact, the evolutionary hypothesis is so comprehensive and explains so many of the observations that it can be considered to have "graduated" from being just a hypothesis to being properly called a theory. Today, astronomers are continuing to refine the details of that theory.

The **solar nebula theory** supposes that planets form in the rotating disks of gas and dust around young stars (■ Figure 8-1). You have seen clear evidence that disks of gas and dust are common around young stars (discussed in Chapter 14). Bipolar flows from protostars were the first evidence of such disks, but modern techniques can image the disks directly (■ Figure 8-2). The evidence is strong that our own planetary system formed in such a

Demodex folliculorum has been found in 97 percent of individuals and is a characteristic of healthy skin.

How Do We Know? 8-1

Two Kinds of Theories: Catastrophic and Evolutionary

How big a role have sudden, catastrophic events played in the history of the solar system? Many theories in science can be classified as either evolutionary, in that they involve gradual processes, or catastrophic, in that they depend on specific, unlikely events. Scientists have generally preferred evolutionary theories. Nevertheless, catastrophic events do occur.

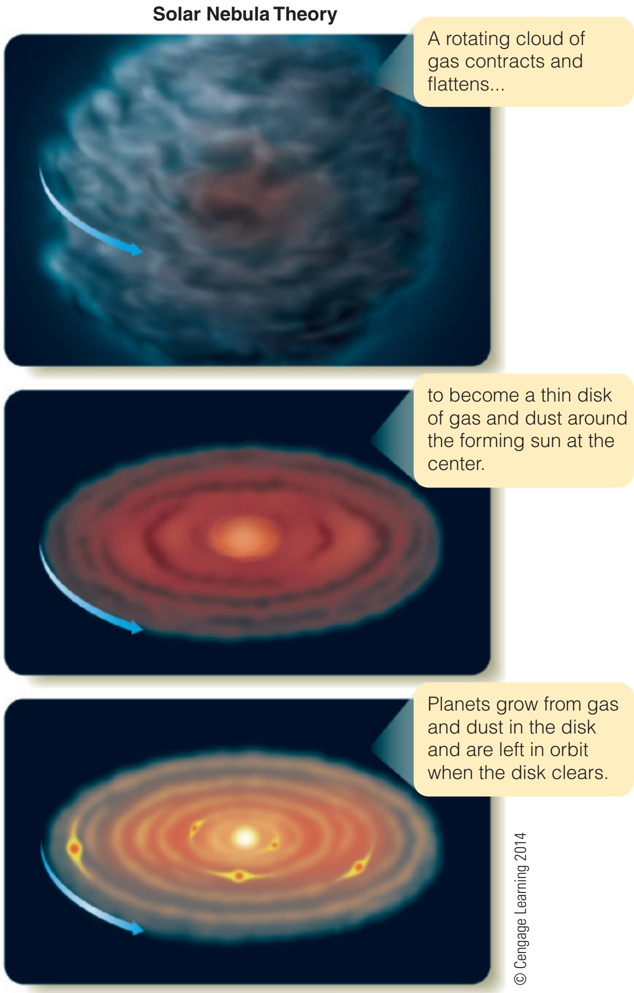
Some people prefer catastrophic theories, perhaps because they like to see spectacular violence from a safe distance, which may explain the success of movies that include lots of car crashes and explosions. Also, catastrophic theories resonate with scriptural accounts of cataclysmic events and special acts of creation. Thus, many people have an interest in catastrophic theories.

Nevertheless, most scientific theories are evolutionary. Such theories do not depend on unlikely events or special acts. For example, geologists study theories of mountain building that are evolutionary and describe mountains being pushed up slowly as millions of years pass. The evidence of erosion and the folded rock layers show that the process is gradual. Because most such natural processes are evolutionary, scientists sometimes find it difficult to accept any theory that depends on catastrophic events.

You will see in this and later chapters that catastrophes do occur. The planets, for example, are bombarded by debris from space, and some of those impacts are very large. As you study astronomy or any other natural science, notice that most theories are evolutionary but that you need to allow for the possibility of unpredictable catastrophic events.


Janet Seeds

Mountains evolve to great heights by rising slowly, not catastrophically.



disk-shaped cloud around the sun. When the sun became luminous enough, the remaining gas and dust were blown away into space, leaving the planets orbiting the sun.

According to the solar nebula theory, Earth and the other planets of the solar system formed billions of years ago as the sun condensed from a cloud of interstellar gas and dust. If planet formation is a natural part of star formation, most stars should have planets.

SCIENTIFIC ARGUMENT

Why does the solar nebula theory imply planets are common? Often, the implications of a theory are more important in building a scientific argument than the theory itself. The solar nebula theory is an evolutionary theory, and if it is correct, the planets of our solar system formed from the disk of gas and dust that surrounded the sun as it condensed from the interstellar medium. That suggests it is a common process. Most stars form with disks of gas and dust around them, and planets should form in such disks. Planets should therefore be very common in the universe.

Now build a new scientific argument. **Why would a catastrophic hypothesis for the formation of the solar system suggest that planets are not common?**

Figure 8-1

The solar nebula theory implies that the planets formed along with the sun.

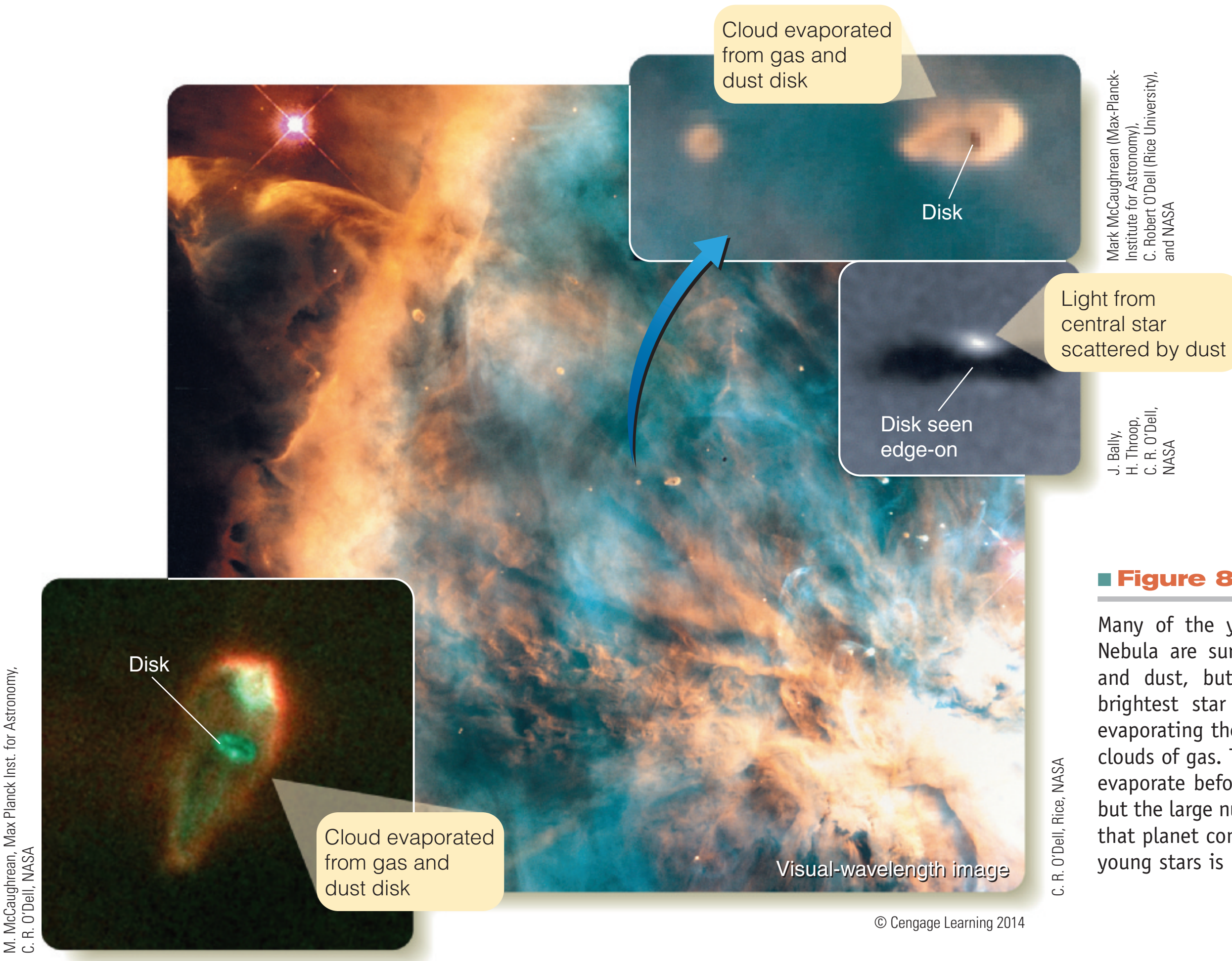


Figure 8-2

Many of the young stars in the Orion Nebula are surrounded by disks of gas and dust, but intense light from the brightest star in the neighborhood is evaporating the disks to form expanding clouds of gas. These particular disks may evaporate before they can form planets, but the large number of such disks shows that planet construction material around young stars is common.

8-2 A Survey of the Solar System

To test their hypotheses about how the solar system was born, astronomers searched the present solar system for evidence of its past. In this section you will survey the solar system and compile a list of its most significant characteristics that are potential clues to how it formed.

You can begin with the most general view of the solar system. It is, in fact, almost entirely empty space (look back to Figure 1-7). Imagine reducing the scale of the solar system until Earth is the size of a grain of table salt, about 0.3 mm (0.01 in.) in diameter. The moon is a speck of pepper about 1 cm (0.4 in.) away, and the sun is the size of a small plum located 4 m (13 ft) from Earth. Jupiter is an apple seed 20 m (66 ft) from the sun. Neptune, at the edge of the planetary zone, is a large grain of sand over 120 m (400 ft) from the central plum. Your model solar system would be larger than two football fields and would need a powerful microscope to detect the asteroids orbiting between Mars and Jupiter. The planets are tiny specks of matter scattered around the sun—the last remains of the solar nebula.

Revolution and Rotation

The planets revolve* around the sun in orbits that lie close to a common plane. The orbit of Mercury, the closest planet to the sun, is tipped 7.0° to Earth's orbit. The rest of the planets' orbital planes are inclined by no more than 3.4°. As you can see, the solar system is basically flat and disk shaped.

The rotation of the sun and planets on their axes also seems related to the rotation of the disk. The sun rotates with its equator inclined only 7.2° to Earth's orbit, and most of the other planets' equators are tipped less than 30°. The rotations of Venus and Uranus are peculiar, however. Venus rotates backward compared with the other planets, whereas Uranus rotates on its side with its equator almost perpendicular to its orbit. You will explore these planets in detail in Chapters 12 and 13, but later in this chapter you will be able to understand how they might have acquired their peculiar rotations.

*Recall from Chapter 2 that the words *revolve* and *rotate* refer to different types of motion. A planet revolves around the sun but rotates on its axis. Cowboys in the Old West didn't carry revolvers. They actually carried rotators. And you don't rotate your tires every 6 months, you actually revolve them.

There is a preferred direction of motion in the solar system—counterclockwise as seen from the north. All the planets revolve around the sun in that direction. With the exception of Venus and Uranus, all the planets also rotate on their axes in that direction. Furthermore, nearly all of the moons in the solar system, including Earth's moon, orbit around their respective planets in that same direction. With only a few exceptions, revolution and rotation in the solar system follow a single theme. Apparently, these motions today are related to the original rotation of a disk of solar system construction material.

Two Kinds of Planets

Perhaps the most striking clue to the origin of the solar system comes from the obvious division of the planets into two groups, the small Earth-like worlds and the giant Jupiter-like worlds. The difference is so dramatic that you are led to say, "Aha, this must mean something!" Study **Terrestrial and Jovian Planets** on pages 146–147, notice three important points, and learn two new terms:

1. The two kinds of planets are distinguished by their location. The four inner *Terrestrial planets* are quite different from the four outer *Jovian planets.*
2. Craters are common. Almost every solid surface in the solar system is covered with craters.
3. The two groups of planets are also distinguished by properties such as number of moons and presence or absence of rings. A theory of the origin of the planets needs to explain those properties.

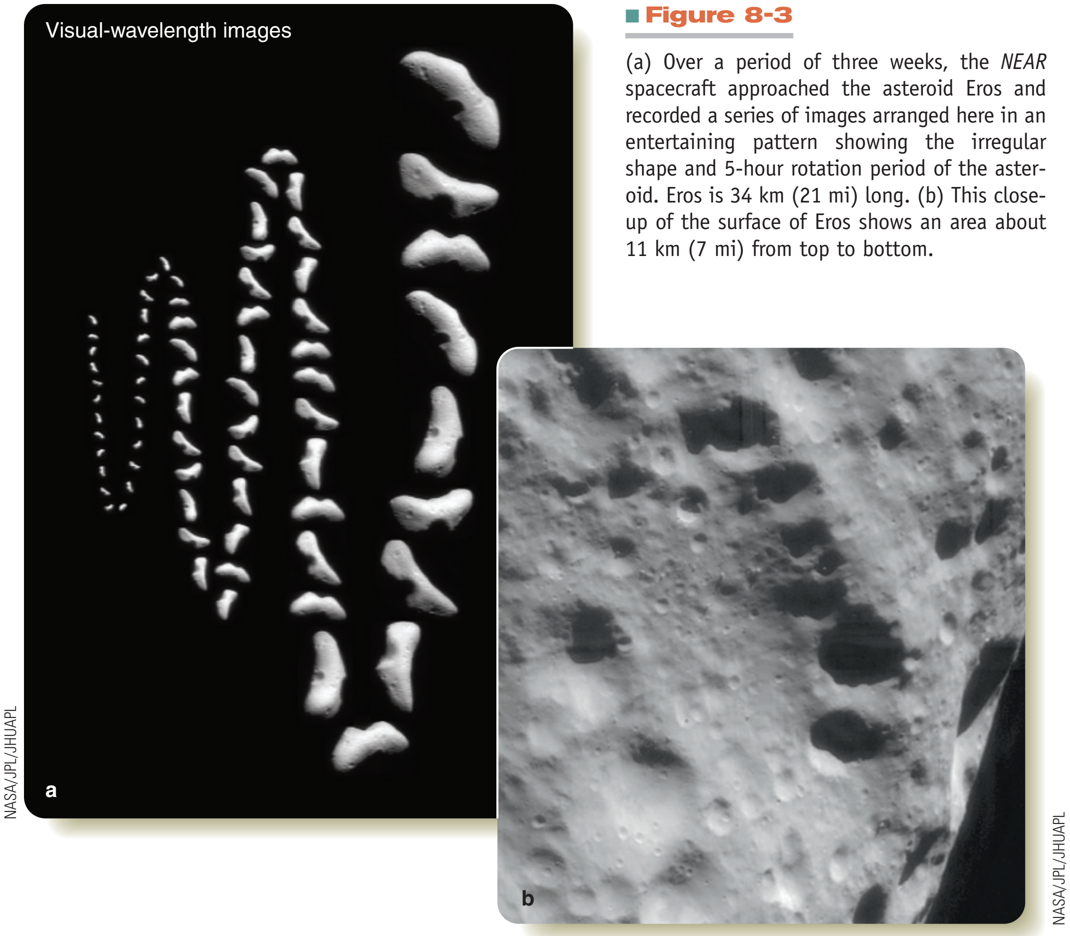
The division of the planets into two groups is a clue to how our solar system formed. The present properties of individual planets, however, don't tell everything you need to know about their origins. The planets have all evolved since they formed. For further clues about the origin of the planets, you can look at smaller objects that have remained largely unchanged since soon after the birth of the solar system.

Cosmic Debris

The sun and planets are not the only products of the solar nebula. The solar system is littered with several kinds of space debris: asteroids, Kuiper belt objects, comets, and meteoroids. These objects will be considered in further detail in Chapter 12, but they are mentioned here because they are a rich source of information about the origin of the planets.

The **asteroids,** sometimes called minor planets, are small rocky worlds, most of which orbit the sun in a belt between the orbits of Mars and Jupiter. More than 100,000 asteroids have orbits that are charted. It is a **Common Misconception** that the asteroids are the remains of a planet that broke apart. In fact, planets are held together very tightly by their gravity and do not "break apart." Astronomers recognize the asteroids as debris left over from the failure of a planet to form at a distance of about 3 AU from the sun.

Spectroscopic observations indicate that asteroid surfaces are made up of a variety of rocky and metallic materials. (Note that *metal* is used here in the familiar sense, referring to substances like iron, rather than the stellar astronomer's sense meaning any element other than hydrogen and helium.) Photographs returned by robotic spacecraft show that asteroids are generally irregular in shape and covered with craters (■ Figure 8-3). Those observations will be discussed in detail in Chapter 12, but in this quick survey of the solar system you have enough information to conclude that the solar nebula included elements that compose rock and

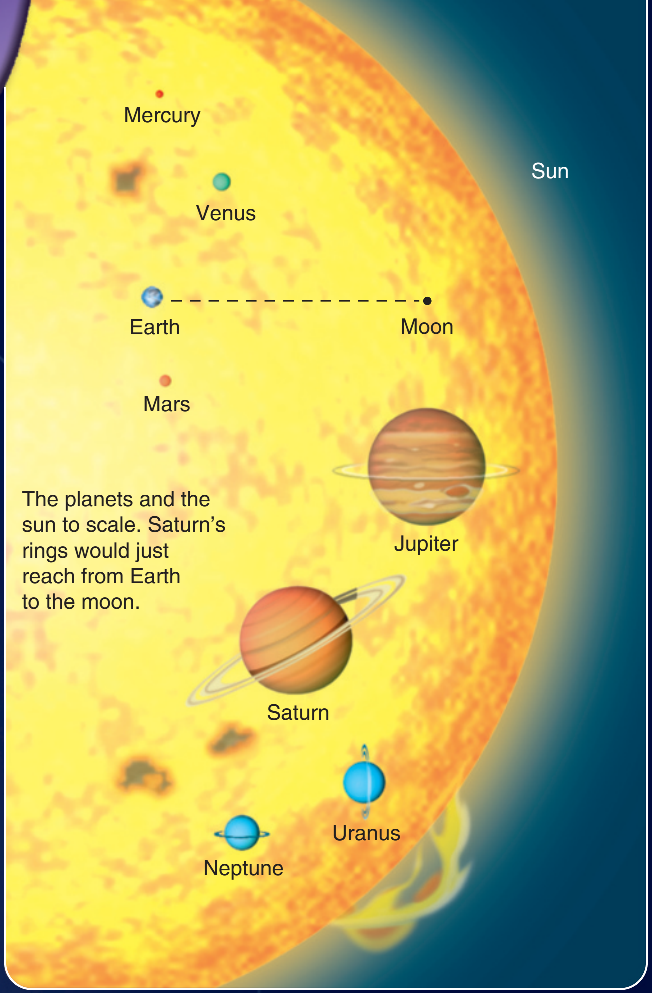


■ **Figure 8-3**

(a) Over a period of three weeks, the *NEAR* spacecraft approached the asteroid Eros and recorded a series of images arranged here in an entertaining pattern showing the irregular shape and 5-hour rotation period of the asteroid. Eros is 34 km (21 mi) long. (b) This close-up of the surface of Eros shows an area about 11 km (7 mi) from top to bottom.

Terrestrial and Jovian Planets

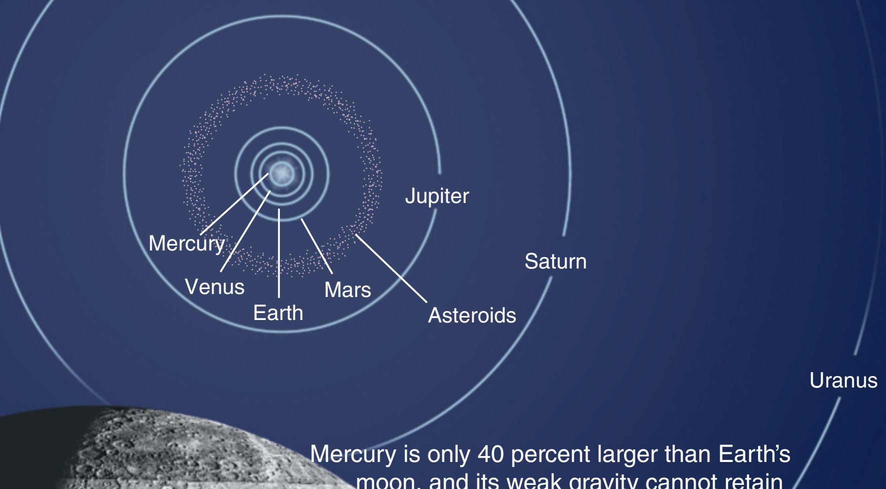
1 The distinction between the Terrestrial planets and the Jovian planets is dramatic. The inner four planets, Mercury, Venus, Earth, and Mars, are **Terrestrial planets,** meaning they are small, dense, rocky worlds with little or no atmosphere. The outer four planets, Jupiter, Saturn, Uranus, and Neptune, are **Jovian planets,** meaning they are large, low-density worlds with thick atmospheres and liquid or ice interiors.



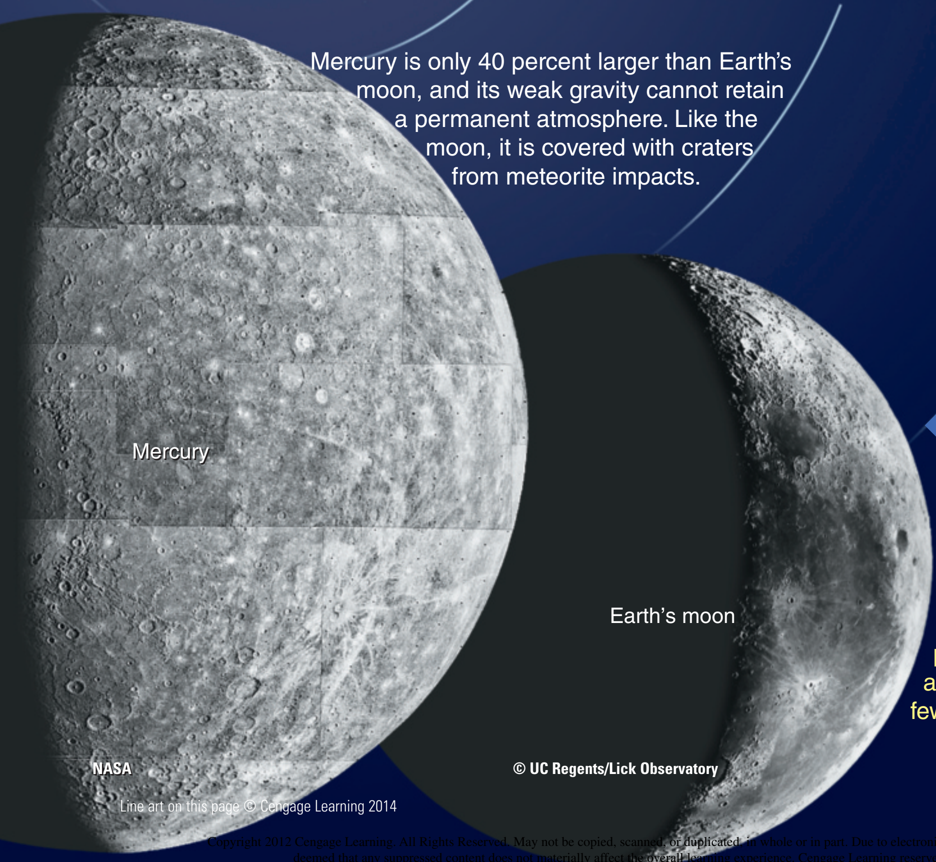
The planets and the sun to scale. Saturn's rings would just reach from Earth to the moon.

1a Of the Terrestrial planets, Earth is most massive, but the Jovian planets are much more massive. Jupiter contains over 300 Earth masses, Saturn nearly 100 Earth masses. Uranus and Neptune, respectively, contain 15 and 17 Earth masses.

Planetary orbits to scale. The Terrestrial planets lie quite close to the sun, whereas the Jovian planets are spread far from the sun.



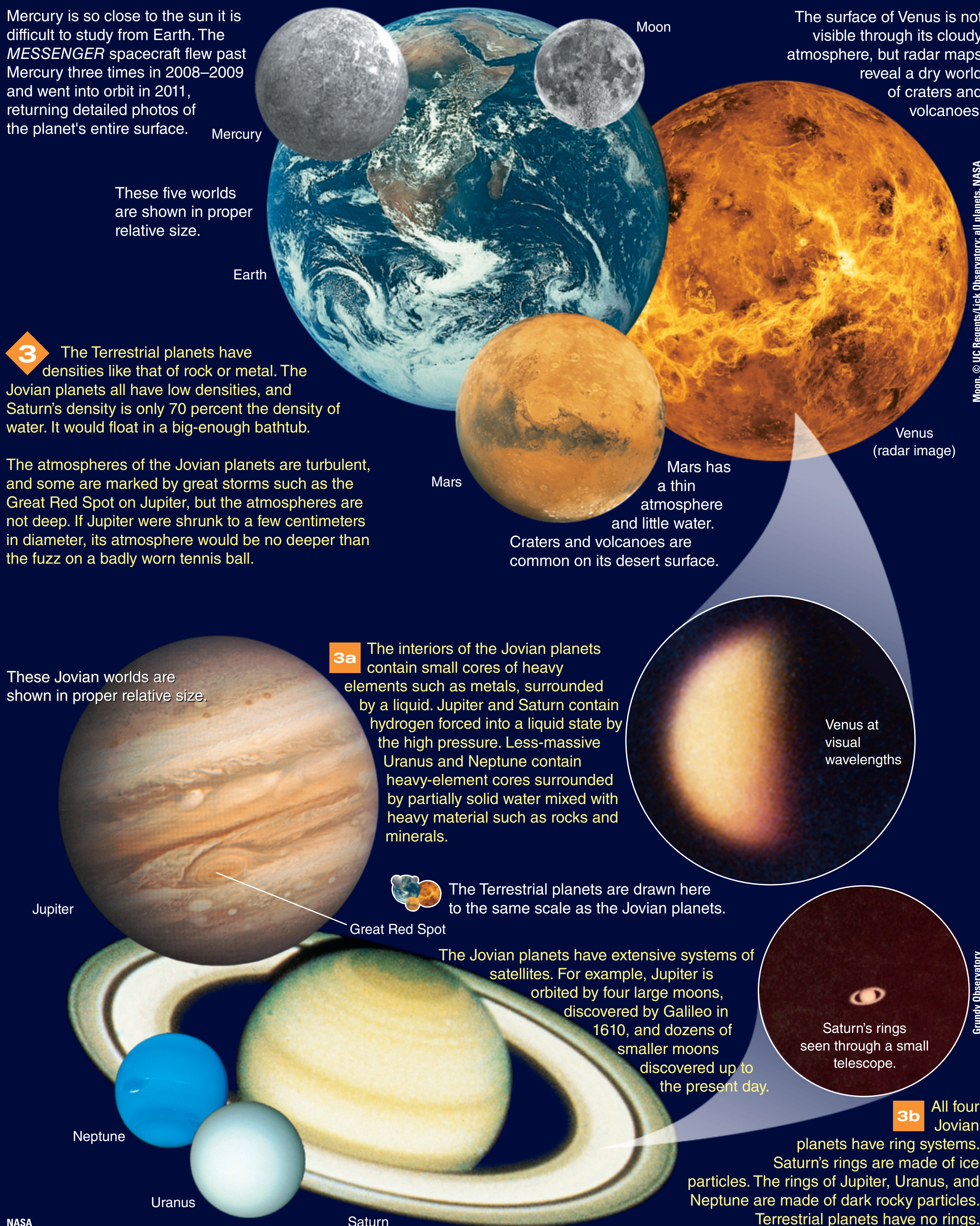
Mercury is only 40 percent larger than Earth's moon, and its weak gravity cannot retain a permanent atmosphere. Like the moon, it is covered with craters from meteorite impacts.



NASA

© UC Regents/Lick Observatory

2 Craters are common on all of the surfaces in the solar system that are strong enough to retain them. Earth has about 150 impact craters, but many more have been erased by erosion. Besides the planets, the asteroids and nearly all of the moons in the solar system are scarred by craters. Ranging from microscopic to hundreds of kilometers in diameter, these craters have been produced over the ages by meteorite impacts. When astronomers see a rocky or icy surface that contains few craters, they know that the surface is young.



Line art on this page © Cengage Learning 2014

metal, and also that collisions have played an important role in the solar system's history.

Since 1992, astronomers have discovered more than a thousand small, dark, icy bodies orbiting in the outer fringes of the solar system beyond Neptune. This collection of objects is called the **Kuiper belt** after astronomer Gerard Kuiper (KYE-per), who predicted their existence in the 1950s. There are probably 100 million objects larger than 1 km, in the Kuiper belt, many more than in the asteroid belt. A successful theory for the formation of the solar system should also explain how the **Kuiper belt objects (KBOs)** came to be where they are. You will find out more about the origin of the Kuiper belt in Chapter 11.

In contrast to the rocky asteroids and dark Kuiper belt objects, the brightest **comets** can be seen with the naked eye and are impressively beautiful objects (■ Figure 8-4). A comet may be visible for many weeks as it sweeps through the inner solar system. Most comets are faint, however, and difficult to locate even at their brightest.

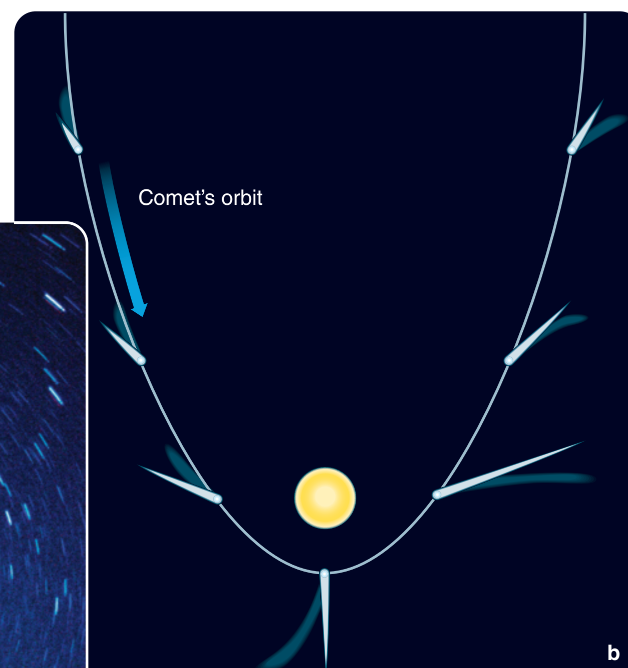
The nuclei of comets are ice-rich bodies a few kilometers or tens of kilometers in diameter, similar in size to asteroids. Like asteroids, comets are understood to be left over from the origin of the planets and provide evidence that at least some parts of the solar nebula had abundant icy material. You will discover more about the composition and history of comets in Chapter 12.

A comet nucleus remains frozen and inactive while it is far from the sun. If the comet's orbit carries it into the inner solar system, the sun's heat begins to vaporize the ices, releasing gas and dust. The flow of solar wind, plus **radiation pressure** exerted by sunlight, pushes the gas and dust away, forming a long tail. As a result the tail of a comet always points approximately away from the sun (Figure 8-4b), no matter what direction the comet itself is moving. The beautiful tail of a comet can be longer than an AU although is produced by a relatively tiny nucleus only a few kilometers in diameter.

Unlike the stately comets, **meteors** flash across the sky in momentary streaks of light (■ Figure 8-5). They are commonly called "shooting stars." Of course, they are not stars but small bits of rock and metal colliding with Earth's atmosphere and bursting into incandescent vapor because of friction with the air about 80 km (50 mi) above the ground. This vapor condenses to form dust that settles slowly to Earth, adding about 40,000 tons per year to the planet's mass.

Technically, the word *meteor* refers to the streak of light in the sky. In space, before its fiery plunge, the object is called a

Kent Wood/Photo Researchers, Inc.



Celestron International

■ Figure 8-4

Comets orbit the sun in long, elliptical orbits and become visible when the sun's heat vaporizes its ices and pushes the gas and dust into a tail pointing away from the sun. (a) A comet may remain visible in the evening or morning sky for weeks as it moves through the inner solar system. Although comets are moving rapidly along their orbits, they are so distant that, on any particular evening, a comet seems to hang motionless in the sky. Comet Hyakutake is shown here near Polaris in 1996. (b) A comet in a long, elliptical orbit becomes visible when the sun's heat vaporizes its ices and pushes the gas and dust away in a tail.

Daniel Good

■ **Figure 8-5**

A meteor is a sudden streak of glowing gases produced by a bit of material falling into Earth's atmosphere. Friction with the air vaporizes the material about 80 km (50 mi) above Earth's surface. This meteor is seen against the background of part of the Milky Way.

meteoroid, and any part of it that survives its fiery passage to Earth's surface is called a **meteorite.** Most meteoroids are specks of dust, grains of sand, or tiny pebbles. Almost all the meteors you see in the sky are produced by meteoroids that weigh less than 1 gram. Only rarely is a meteoroid massive enough and strong enough to survive its plunge and reach Earth's surface.

Thousands of meteorites have been found, and you will learn more about their various types in Chapter 12. Meteorites are mentioned here for one specific clue they can give you concerning the solar nebula: Meteorites can reveal the age of the solar system.

Age of the Solar System

According to the solar nebula theory, the planets should be about the same age as the sun. The most accurate way to find the age of a rocky body is to bring a sample into the laboratory and analyze the radioactive elements it contains (**How Do We Know 8-2**).

When a rock solidifies, it incorporates known percentages of the chemical elements. A few of these elements have forms called isotopes (look back to Chapter 6) that are radioactive, meaning they gradually decay into other isotopes. For example, potassium-40, called a parent isotope, decays into calcium-40 and argon-40, called daughter isotopes. The **half-life** of a radioactive substance is the time it takes for half of the parent isotope atoms to decay into daughter isotope atoms. The abundance of a radioactive substance gradually decreases as it decays, and the abundances of the daughter substances gradually increase (■ Figure 8-6). The half-life of potassium-40 is 1.3 billion years. If you also have information about the abundances of the elements in the original rock, you can compare those with the present abundances and find the age of the rock. For example, if you study a rock and find that only 50 percent of the potassium-40 remains and the rest has become a mixture of daughter isotopes, you could conclude that one half-life must have passed and that the rock is 1.3 billion years old.

Potassium isn't the only radioactive element used in radioactive dating. Uranium-238 decays with a half-life of 4.5 billion years to form lead-206 and other isotopes. Rubidium-87 decays into strontium-87 with a half-life of 47 billion years. Any of these substances can be used as a radioactive clock to find the age of mineral samples.

Of course, to find a radioactive age, you need to get a sample into the laboratory, and the only celestial bodies of which scientists have samples are Earth, the moon, Mars, and meteorites. The oldest Earth rocks so far discovered and dated are tiny zircon crystals from Australia that are 4.4 billion years old. That does not mean that Earth formed 4.4 billion years ago. As you will see in the next chapter, the surface of Earth is active, and the crust is continually destroyed and reformed with material welling up from beneath the crust. Those types of processes tend to dilute the daughter atoms and spread them away from the parent atoms, effectively causing the radioactive clocks to reset to zero. The radioactive age of a rock is actually the length of time since the material in that rock was last melted. Consequently, the dates of these oldest rocks tells you only a lower limit to the age of Earth, in other words, that Earth is *at least* 4.4 billion years old.

One of the most exciting goals of the *Apollo* lunar landings was to bring lunar rocks back to Earth's laboratories where their ages could be measured. Because the moon's surface is not geologically active like Earth's surface, some moon rocks might have survived unaltered since early in the history of the solar system. In fact, the oldest moon rocks are 4.5 billion years old. That means the moon must be *at least* 4.5 billion years old.

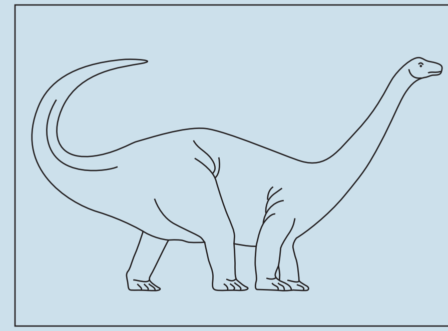
Although no one has yet been to Mars, over a dozen meteorites found on Earth have been identified by their chemical composition as having come from Mars. Most of these have ages of only a billion years or so, but one has an age of approximately 4.5 billion years. Mars must be at least that old.

Reconstructing the Past from Evidence and Hypotheses

How can we know how the solar system formed? Scientists often solve problems in which they must reconstruct the past. Some of these reconstructions are obvious, such as an archaeologist excavating the ruins of a burial tomb, but others are less obvious. In each case, success requires the interplay of hypotheses and evidence to re-create a past that no longer exists.

It is obvious that astronomers reconstruct the past when they use evidence gathered from meteorites to study the origin of the solar system, but a biologist studying a centipede is also reconstructing the past. How did this creature come to have a segmented body with so many legs? How did it evolve the metabolism that allows it to move quickly and hunt prey? Although the problem might at first seem to be one of mere anatomy, the scientist must reconstruct an environment that no longer exists.

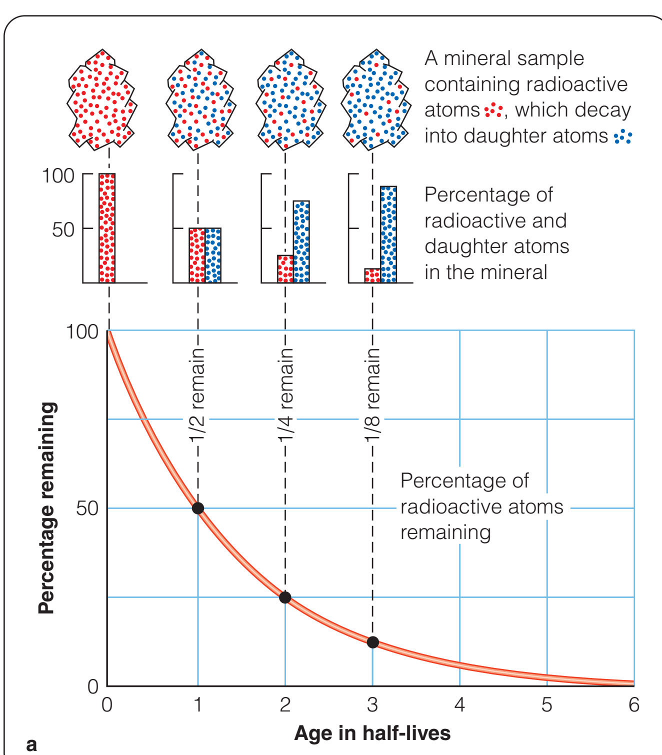
The astronomer's problem is not just to understand what the planets are like but also to understand how they got that way. That means planetary scientists must look at the evidence they can see today and reconstruct a history of the solar system, a past that is quite different from the present. If you had a time machine, it would be a fantastic adventure to go back and watch the planets form. Time machines are impossible, but scientists can use the scientific method's grand interplay of evidence and hypotheses to journey back billions of years and reconstruct a past that no longer exists.



© Cengage Learning 2014

One way science can enrich and inform our lives is by re-creating a world that no longer exists.

The most important source for determining the age of the solar system is meteorites. Radioactive dating of meteorites yields a range of ages, but there is a fairly precise upper limit—many meteorite samples have ages of 4.56 billion years old, and none are older. That figure is widely accepted as the age of the solar system and is often rounded to 4.6 billion years. The true ages of Earth, the moon, and Mars are also assumed to be 4.6 billion years, although no rocks from those bodies have yet been found that have remained unaltered for that entire stretch of time.



© Cengage Learning 2014

Lab photograph, R. Kempton, New England Meteoritical

Figure 8-6

(a) The radioactive atoms (red) in a mineral sample decay into daughter atoms (blue). Half the radioactive atoms are left after one half-life, a fourth after two half-lives, an eighth after three half-lives, and so on. (b) Radioactive dating shows that this fragment of the Allende meteorite is 4.56 billion years old. It contains a few even older interstellar grains, which formed long before our solar system.

Table 8-1 | Characteristic Properties of the Solar System

1. Disk shape of the solar system
   Orbits in nearly the same plane
   Common direction of rotation and revolution
2. Two planetary types
   Terrestrial—inner planets; high density
   Jovian—outer planets; low density
3. Planetary rings and large satellite systems
   YES for Jupiter, Saturn, Uranus, and Neptune
   NO for Mercury, Venus, Earth, and Mars
4. Space debris—asteroids, comets, and meteoroids
   Composition, Orbits
   Asteroids in inner solar system, composition like Terrestrial planets
   Comets in outer solar system, composition like Jovian planets
5. Common age of about 4.6 billion years measured or inferred for Earth, the moon, Mars, meteorites, and the sun

One last celestial body deserves mention: the sun. Astronomers estimate the age of the sun to be about 5 billion years, but that is not a radioactive date because we have no samples of radioactive material from the sun. Instead, an independent estimate for the age of the sun can be made using helioseismological observations and mathematical models of the sun's interior (look back to Chapter 7 and ahead to Chapter 14). This yields a value of about 5 billion years, plus or minus 1.5 billion years, a number that is in agreement with the age of the solar system derived from the age of meteorites.

Apparently, all the bodies of the solar system formed at about the same time, some 4.6 billion years ago. You can add this as the final item to your list of characteristic properties of the solar system (■ Table 8-1).

SCIENTIFIC ARGUMENT

In what ways does the solar system resemble a disk?

Notice that this argument is really a summary of pieces of evidence. First, the general shape of the solar system is that of a disk. The orbit of Mercury is inclined 7.0° to the plane of Earth's orbit, and the rest of the planets are in orbits inclined less than that. In other words, the planets are confined to a thin disk with the sun at its center.

Second, the motions of the sun and planets also follow this disk theme. The sun and most of the planets rotate in the same direction, with their equators near the plane of the solar system. Also, all of the planets revolve around the sun in that same direction. The objects in our solar system mostly move in the same direction, which further reflects a disk theme.

One of the basic characteristics of our solar system is its disk shape, but another dramatic characteristic is the division of the planets into two groups. Build an argument to detail that evidence. **What are the distinguishing differences between the Terrestrial and Jovian planets?**

8-3 The Story of Planet Building

The challenge for modern planetary scientists is to compare the characteristics of the solar system with predictions of the solar nebula theory so they can work out details of how the planets formed.

Chemical Composition of the Solar Nebula

Everything astronomers know about the solar system and star formation suggests that the solar nebula was a fragment of an interstellar gas cloud. Such a cloud would have been mostly hydrogen with some helium and small amounts of the heavier elements.

That is precisely what you see in the composition of the sun (look back at Table 7-1). Analysis of the solar spectrum shows that the sun is mostly hydrogen, with a quarter of its mass being helium and only about 2 percent being heavier elements. Of course, nuclear reactions have fused some hydrogen into helium, but this happens in the sun's core and has not affected the composition of its surface and atmosphere, which are the parts you can observe directly. That means the composition revealed in the sun's spectrum is essentially the composition of the gases from which the sun formed.

This must have been the composition of the solar nebula, and you can also see that composition reflected in the chemical compositions of the planets. The inner planets are composed of rock and metal, and the outer planets are rich in low-density gases such as hydrogen and helium. The chemical composition of Jupiter resembles the composition of the sun, but if you allowed low-density gases to escape from a blob of stuff with the same overall composition as the sun or Jupiter, the relative proportions of the remaining heavier elements would resemble the chemical composition of Earth and the other Terrestrial planets.

Condensation of Solids

An important clue to understanding the process that converted the nebular gas into solid matter is the variation in density among solar system objects. You have already noted that the four inner planets are small and have high densities, resembling Earth, whereas the outermost planets are large and have low density, resembling Jupiter.

Even among the four Terrestrial planets, you will find a pattern of slight differences in density. Merely listing the observed densities of the Terrestrial planets does not reveal the pattern clearly because Earth and Venus, being more massive, have stronger gravity and have squeezed their interiors to higher densities. The **uncompressed densities**—the densities the planets would have if their gravity did not compress them, or to put it another way, the average densities of their original construction materials—can be calculated using

Table 8-2 | Observed and Uncompressed Densities

Planet	Observed Density (g/cm^3)	Uncompressed Density (g/cm^3)
Mercury	5.44	5.30
Venus	5.24	3.96
Earth	5.50	4.07
Mars	3.94	3.73

Table 8-3 | The Condensation Sequence

Temperature (K)	Condensate	Planet (Estimated Temperature of Formation; K)
1500	Metal oxides	Mercury (1400)
1300	Metallic iron and nickel	
1200	Silicates	
1000	Feldspars	Venus (900)
680	Troilite (FeS)	Earth (600)
		Mars (450)
175	H_2O ice	Jovian (175)
150	Ammonia-water ice	
120	Methane-water ice	
65	Argon-neon ice	Pluto (65)

the actual densities and masses of each planet (Table 8-2). In general, the closer a planet is to the sun, the higher its uncompressed density.

According to the solar nebula theory, the observed pattern of planet densities originated when solid grains first formed, a process called **condensation.** Solid particles condensed from the gas of the nebula as it cooled. The kind of matter that could condense in a particular region depended on the temperature of the gas there. In the inner regions, close to the sun, the temperature was evidently 1500 K or so. The only materials that can form grains at that temperature are compounds with high melting points, such as metal oxides and pure metals, which are very dense. Farther out in the nebula it was cooler, and silicates (rocky material) could also condense, in addition to metal. These are less dense than metal oxides and metals. Mercury, Venus, Earth, and Mars are evidently composed of a mixture of metals, metal oxides, and silicates, with proportionately more metals close to the sun and more silicates farther from the sun.

Even farther from the sun there was a boundary called the **ice line** beyond which water vapor could freeze to form ice particles. Yet a little farther from the sun, compounds such as methane and ammonia could condense to form other types of ice. Water vapor, methane, and ammonia were abundant in the solar nebula, so beyond the ice line the nebula would have been filled with a blizzard of ice particles, mixed with small amounts of silicate and metal particles that could also condense there. Those ices are low-density materials. The densities of Jupiter and the other outer planets correspond to a mix of ices plus relatively small amounts of silicates and metal.

The sequence in which the different materials condense from the gas as you move away from the sun toward lower temperature is called the **condensation sequence** (Table 8-3). It suggests that the planets, forming at different distances from the sun, should have accumulated from different kinds of materials in a predictable way.

People who have read a little bit about the origin of the solar system may hold the **Common Misconception** that the matter in the solar nebula was sorted by density, with the heavy rock and metal sinking toward the sun and the low-density gases being blown outward. That is not the case. The chemical composition of the solar nebula was originally roughly the same throughout the disk. The important factor was temperature: The inner nebula was hot, and only metals and rock could condense there, whereas the cold outer nebula could form lots of ices along with metals and rock. The ice line seems to have been between Mars and Jupiter, and it separates the region for formation of the high-density Terrestrial planets from that of the low-density Jovian planets.

Formation of Planetesimals

In the development of the planets from the material of the solar nebula disk, three processes operated to collect solid bits of matter—metal, rock, ice—into larger bodies called **planetesimals,** which eventually made the planets. The study of planet building is the study of these three processes.

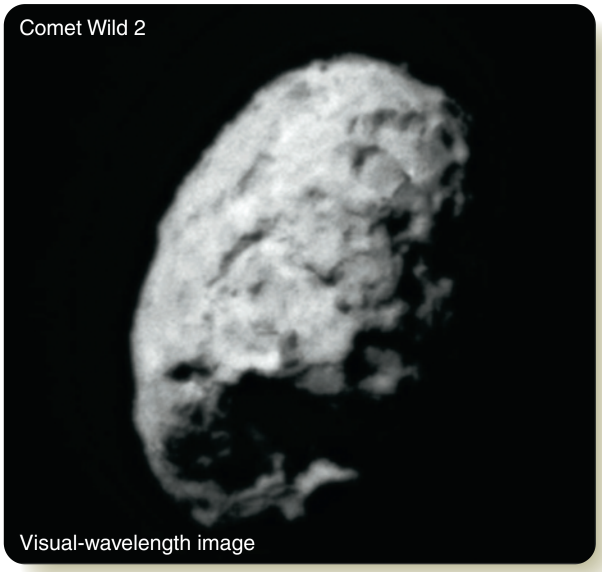
In the previous section you learned a little about the process of condensation. Planetary development in the solar nebula began with the formation of dust grains by condensation. A particle grows by condensation when it adds matter one atom or molecule at a time from a surrounding gas. Snowflakes, for example, grow by condensation in Earth's atmosphere. In the solar nebula, dust grains would have been bombarded continuously by atoms of gas, and some of those stuck to the grains. A microscopic grain capturing a layer of gas molecules on its surface increases its mass by a much larger fraction than a gigantic boulder capturing a single layer of molecules. For that reason condensation can increase the mass of a small grain rapidly, but, as the grain grows larger, condensation becomes less effective, and other processes became more important.

The second process of planetesimal formation is **accretion,** the sticking together of solid particles. You may have seen accretion in action if you have walked through a snowstorm with big,

fluffy flakes. If you caught one of those “flakes” on your glove and looked closely, you saw that it was actually made up of many tiny, individual flakes that had collided as they fell and accreted to form larger particles. In the solar nebula, the dust grains were, on average, no more than a few centimeters apart, so they collided frequently and could accrete into larger particles.

When the particles grew to sizes larger than a centimeter, they would have been subject to new processes that tended to concentrate them. One important effect was that the growing solid objects would have collected into the plane of the solar nebula. Small dust grains could not fall into the plane because the turbulent motions of the gas kept them stirred up, but larger objects had more mass, and gas motions could not have prevented them from settling into the plane of the spinning nebula. Astronomers calculate this would have concentrated the larger solid particles into a relatively thin layer about 0.01 AU thick that would have made further growth more rapid. There is no clear distinction between a very large grain and a very small planetesimal, but you might consider an object to be a planetesimal when its diameter approaches a kilometer (0.6 mi) or so (■ Figure 8-7).

This concentration of large particles and planetesimals into the plane of the nebula is analogous to the flattening of a forming galaxy, and a process also found in galaxies may have become important once the plane of planetesimals formed. Computer models show that the rotating disk of particles should have been gravitationally unstable and would have been disturbed by spiral density waves resembling the much larger ones found in spiral galaxies. Those waves could have further concentrated the planetesimals and helped them coalesce into objects up to 100 km (60 mi) in diameter.



■ **Figure 8-7**

What did the planetesimals look like? You can get a clue from this photo of the 5-km-wide nucleus of Comet Wild 2 (pronounced *Vildt-two*). Whether rocky or icy, the planetesimals must have been small, irregular bodies, scarred by craters from collisions with other planetesimals.

Through these processes, according to the solar nebula theory, the disk of gas and dust around the forming sun became filled with trillions of solid particles ranging in size from pebbles to tiny planets. As the largest began to exceed 100 km in diameter, a third process began to affect them, and a new stage in planet building began, the formation of protoplanets.

Growth of Protoplanets

The coalescing of planetesimals eventually formed **protoplanets,** the name for massive objects destined to become planets. As these larger bodies grew, a new process helped them grow faster and altered their physical structure.

If planetesimals had collided at orbital velocities, they would have been unable to stick together. A typical orbital velocity in the solar system is about 10 km/s (22,000 mph). Head-on collisions at this velocity would vaporize the material. However, the planetesimals were all moving in the same direction in the nebular plane and didn’t collide head on. Instead, they merely “rubbed shoulders,” so to speak, at low relative velocities. Such gentle collisions would have been more likely to combine planetesimals than to shatter them.

The largest planetesimals would grow the fastest because they had the strongest gravitational field. Their stronger gravity could attract additional material. Their gravity could also hold on to a cushioning layer of soil that would trap fragments. Astronomers calculate that the largest planetesimals would have grown quickly to protoplanetary dimensions, sweeping up more and more material.

Protoplanets had to begin growing by accumulating solid material because they did not have enough gravity to capture and hold large amounts of gas. In the warm solar nebula, the atoms and molecules of gas were traveling at velocities much larger than the escape velocities of modest-sized protoplanets. Therefore, in their early development, the protoplanets could grow only by attracting solid bits of rock, metal, and ice. Once a protoplanet approached a size of 15 Earth masses or so, however, it could begin to grow by **gravitational collapse,** the rapid accumulation of large amounts of infalling gas from the nebula.

The theory of protoplanet growth into planets supposes that all the planetesimals had about the same chemical composition. The planetesimals accumulated to form a planet-sized ball of material with homogeneous composition throughout. Once the planet formed, heat would begin to accumulate in its interior from the decay of short-lived radioactive elements.

The violent impacts of infalling particles would also have released energy called **heat of formation.** These two heating sources would eventually have melted the planet and allowed it to differentiate. **Differentiation** is the separation of material according to density. Once a planet melted, the heavy metals such as iron and nickel, plus elements chemically attracted to them, would settle to the core, while the lighter silicates and related materials floated to the surface to form a low-density

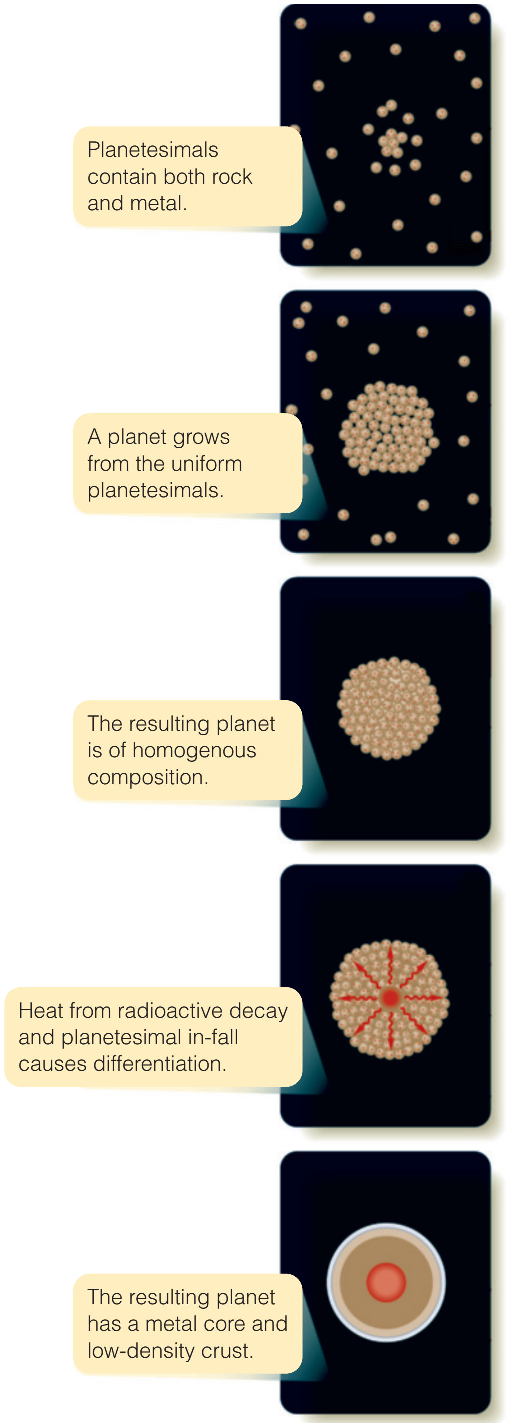
crust. The scenario of planetesimals combining into planets that subsequently differentiated is shown in ■ Figure 8-8.

The process of differentiation depends partly on the presence of short-lived radioactive elements whose rapid decay would have released enough heat to melt the interior of planets. Astronomers know such radioactive elements were present because very old rock from meteorites contains daughter isotopes such as magnesium-26. That isotope is produced by the decay of aluminum-26 with a half-life of only 0.74 million years. The aluminum-26 and similar short-lived radioactive isotopes are gone now, but they must have been produced in a supernova explosion that occurred shortly before the formation of the solar nebula. In fact, some astronomers suspect that a supernova explosion could have triggered the formation of the sun and other stars by compressing interstellar clouds (to be discussed in Chapter 14). Thus, our solar system may exist because of a supernova explosion that occurred about 4.6 billion years ago.

If planets formed by accretion of planetesimals and were later melted by radioactive decay and heat of formation, then Earth's early atmosphere may have consisted of a combination of gases delivered by planetesimal impacts and released from the planet's interior during differentiation. The creation of a planetary atmosphere from a planet's interior is called **outgassing.** Given the location of Earth in the solar nebula, planetary scientists calculate that gases released from its interior during differentiation would not have included as much water as Earth now has. So, astronomers hypothesize that some of Earth's water and atmosphere accumulated late in the formation of the planet as Earth swept up volatile-rich planetesimals. These icy planetesimals would have formed in the cool outer parts of the solar nebula and could have been scattered toward the Terrestrial planets by encounters with the Jovian planets, creating a comet bombardment.

According to the solar nebula theory, the Jovian planets could begin growing by the same processes that built the Terrestrial planets. However, in the inner solar nebula, only metals and silicates could form solids, so the Terrestrial planets grew slowly. In contrast, the outer solar nebula contained not just solid bits of metals and silicates but also ices that included plentiful hydrogen. Astronomers calculate that the Jovian planets would have grown faster than the Terrestrial planets and quickly become massive enough to begin even faster growth by gravitational collapse, drawing in large amounts of gas from the solar nebula. The Terrestrial planet zone did not include ice particles, so those planets developed relatively slowly and never became massive enough to grow further by gravitational collapse.

The Jovian planets must have reached their present size in no more than about 10 million years, before the sun become hot and luminous enough to blow away the remaining gas in the solar nebula, removing the raw material for further Jovian growth. As you will learn in the next section, disturbances from outside the forming solar system may have reduced the time available for Jovian planet formation even more severely.



■ **Figure 8-8**

This simple model of planet building assumes planets formed from accretion and collision of planetesimals that were of uniform composition, containing both metals and rocky material, and that the planets later differentiated, meaning they melted and separated into layers by density and composition.

The Terrestrial planets, in comparison, grew from solids and not from the gas, so they could have continued to grow by accretion from solid debris left behind after the gas was removed. Mathematical models indicate that the Terrestrial planets were at least half finished within 10 million years but could have continued to grow for another 20 million years or so.

The solar nebula theory has been very successful in explaining the formation of the solar system. But there are some problems, and the Jovian planets are the troublemakers.

The Jovian Problem

New information about the star formation process makes it hard to explain the formation of the Jovian parents, and this has caused astronomers to expand and revise the theory of planet formation (■ Figure 8-9).

The new information is that gas and dust disks around newborn stars don't last long. You have seen images of dusty gas disks around the young stars in the Orion star-forming region (Figure 8-2). Those disks are being evaporated by intense ultraviolet radiation from hot O and B stars forming nearby. Astronomers have calculated that nearly all stars form in clusters containing O and B stars, so this evaporation may happen to most disks. Even if a disk did not evaporate quickly, the gravitational influence of the crowded stars in a cluster could strip away the outer parts of the disk. Those are troublesome observations because they seem to indicate that disks can't last longer than about 10 million years, and many evaporate within the astronomically very short span of 100,000 years or so. That's not long enough to grow a Jovian planet by the combination of condensation, accretion, and gravitational collapse proposed in the standard solar nebular theory.

Yet, Jovian planets are common. In the final section of this chapter, you will see evidence that astronomers have found planets orbiting other stars, and almost all of the planets discovered so far have the mass of Jovian planets. There may also be many Terrestrial planets orbiting those stars that are too small to be detected at present, but the important point is that there are lots of Jovian planets around. How they can form quickly enough, before the disks of raw material evaporate, is referred to as the **Jovian problem.**

Mathematical models of the solar nebula have been produced using specially built computers running programs that take weeks to finish a calculation. The results show that the rotating gas and dust of the solar nebula could have become unstable and formed Jovian planets by skipping straight to the step of gravitational collapse. That is, massive planets may have been able to form by **direct collapse** of gas without first forming a dense core by condensation and accretion of solid material. Jupiters and Saturns can form in these direct collapse models in only a few thousand years. If the Jovian planets formed in this way, they could have formed before the solar nebula disappeared, even if the nebula was eroded quickly by neighboring massive hot stars.

The original solar nebula theory hypothesized that the planets formed by accreting a core and then, if they became massive enough, accelerated growth by gravitational collapse. The proposed modification to the theory suggests that the outer planets could have skipped the core accretion phase.

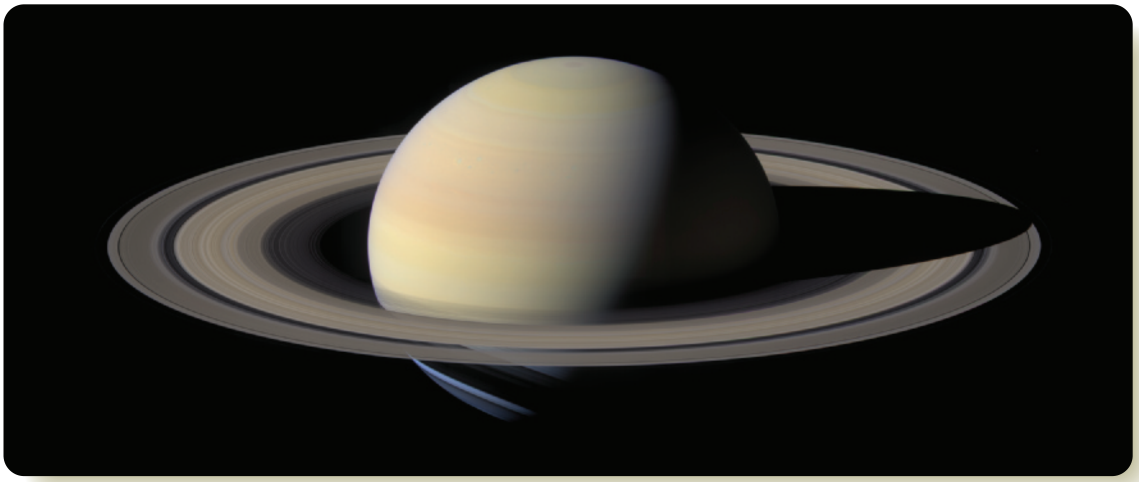
Explaining the Characteristics of the Solar System

Now you have learned enough to put all the pieces of the puzzle together and explain the distinguishing characteristics of the solar system in Table 8-1.

The disk shape of the solar system is inherited from the motion of material in the solar nebula. The sun and planets and moons mostly revolve and rotate in the same direction because

■ **Figure 8-9**

The Jovian worlds pose a problem for modern astronomers. Planet-forming nebulae are blown away in only a few million years by nearby luminous stars, so Jovian planets must form more quickly than initial calculations predicted. Newer research suggests that accretion followed by gravitational collapse can build Jovian planets in about a million years. Under certain conditions, direct gravitational collapse may form some large planets in just thousands of years.



NASA/JPL/Space Science Institute

they formed from the same rotating gas cloud. The orbits of the planets lie in the same plane because the rotating solar nebula collapsed into a disk, and the planets formed in that disk.

The solar nebula hypothesis is evolutionary in that it calls on continuing processes to gradually build the planets. To explain the odd rotations of Venus and Uranus, however, you may need to consider catastrophic events. Uranus rotates on its side. This might have been caused by an off-center collision with a massive planetesimal when the planet was nearly formed, or by gravitational interactions with Saturn and Jupiter. Two hypotheses have been proposed to explain the backward rotation of Venus. Theoretical models suggest that the sun produced tides in the thick atmosphere of Venus that could have eventually reversed the planet's rotation—an evolutionary hypothesis. It is also possible that the rotation of Venus was altered by an off-center impact late in the planet's formation, and that is a catastrophic hypothesis. Both may be true.

The second item in Table 8-1, the division of the planets into Terrestrial and Jovian worlds, can be understood through the condensation sequence. The Terrestrial planets formed in the inner part of the solar nebula, where the temperature was high and only substances such as silicates and metals could condense to form solid particles. That produced the small, dense Terrestrial planets. In contrast, the Jovian planets formed in the outer solar nebula, where the lower temperature allowed the gas to form large amounts of ices. That allowed the Jovian planets to grow rapidly and became massive, low-density worlds. Also, Jupiter and Saturn are so massive they were able to grow by drawing in cool gas by gravitational collapse from the solar nebula. The Terrestrial planets could not do this because they never became massive enough.

The heat of formation (the energy released by infalling matter) was tremendous for these massive planets. Jupiter must have grown hot enough to glow with a luminosity of about 1 percent that of the present sun, although it never got hot enough to generate nuclear energy as a star would. Nevertheless, Jupiter is still hot inside. In fact, both Jupiter and Saturn radiate more heat than they absorb from the sun, so they are evidently still cooling.

A glance at the solar system suggests that you should expect to find a planet between Mars and Jupiter at the present location of the asteroid belt. Mathematical models indicate that the reason asteroids are there rather than a planet is that Jupiter grew into such a massive body that it was able to gravitationally disturb the motion of nearby planetesimals. The bodies that could have formed a planet just inward from Jupiter's orbit instead collided at high speeds and shattered rather than combining, were thrown into the sun, or were ejected from the solar system. The asteroids seen today are the last remains of those rocky planetesimals.

The comets, in contrast, are evidently the last of the icy planetesimals. Some may have formed in the outer solar nebula beyond Neptune, but many probably formed among the Jovian planets where ices could condense easily. Mathematical models show that the massive Jovian planets could have ejected some of these icy planetesimals into the far outer solar system. In a later chapter, you will see evidence that some comets are icy bodies coming from those distant locations, falling back into the inner solar system.

The icy Kuiper belt objects appear to be ancient planetesimals that formed in the outer solar system but were never incorporated into a planet. They orbit slowly far from the light and warmth of the sun and, except for occasional collisions, have not changed much since the solar system was young. The gravitational influence of the planets deflects Kuiper belt objects into the inner solar system where they also are seen as comets.

The large satellite systems of the Jovian worlds may contain two kinds of moons. Some moons may have formed in orbit around forming planets in a miniature version of the solar nebula. Some of the smaller moons, in contrast, may be captured planetesimals, asteroids, and comets. The large masses of the Jovian planets would have made it easier for them to capture satellites.

In Table 8-1, you noted that all four Jovian worlds have ring systems, and you can understand this by considering the large mass of these worlds and their remote location in the solar system. A large mass makes it easier for a planet to hold onto orbiting ring particles; and, being farther from the sun, the ring particles are not as quickly swept away by radiation pressure and the solar wind. It is hardly surprising, then, that the Terrestrial planets, low-mass worlds located near the sun, have no planetary rings.

The last entry in Table 8-1 is the common ages of solar system bodies, and the solar nebula theory has no difficulty explaining that characteristic. According to the theory, the planets formed at the same time as the sun and should have roughly the same age.

Clearing the Nebula

The sun probably formed along with many other stars in a cloud of interstellar material. You have already learned that observations of young stars suggest that radiation and gravitational effects from nearby stars would have tended to disturb and remove the disk of planet construction material around the sun. Even without those external effects, four internal processes would have gradually destroyed the solar nebula.

The most important of these internal processes was radiation pressure. When the sun became a luminous object, light streaming from its photosphere pushed against the particles of the solar nebula. Large bits of matter like planetesimals and planets were not affected, but low-mass specks of dust and individual atoms and molecules were pushed outward and eventually driven from the system.

The second effect that helped clear the nebula was the solar wind, the flow of ionized hydrogen and other atoms away from the sun's upper atmosphere (look back to Chapter 7). This flow is a steady breeze that rushes past Earth at about 400 km/s (250 mi/s). Young stars have even stronger winds than stars of the sun's age and also irregular fluctuations in luminosity, like those observed in young stars such as T Tauri stars, which can accelerate the wind. The strong surging wind from the young sun may have helped push dust and gas out of the nebula.

The third effect that helped clear the nebula was the sweeping up of space debris by the planets. All of the old, solid surfaces in the solar system are heavily cratered by meteorite impacts (■ Figure 8-10). Earth's moon, Mercury, Venus, Mars, and most of the moons in the solar system are covered with craters. A few of these craters have been formed recently by the steady rain of meteorites that falls on all the planets in the solar system, but most of the craters appear to have been formed before roughly 4 billion years ago in what is called the **heavy bombardment,** as the last of the debris in the solar nebula were swept up by the planets.

The fourth effect was the ejection of material from the solar system by close encounters with planets. If a small object such as a planetesimal passes close to a planet, the small object's path will be affected by the planet's gravitational field. In some cases, the small object can gain energy from the planet's motion and be thrown out of the solar system. Ejection is most probable in encounters with massive planets, so the Jovian planets were probably very efficient at ejecting the icy planetesimals that formed in their region of the nebula.

Attacked by the radiation and gravity of nearby stars and racked by internal processes, the solar nebula could not survive very long. Once the gas and dust were gone and most of the planetesimals were swept up, the planets could no longer gain significant mass, and the era of planet building ended.

SCIENTIFIC ARGUMENT

Why are the Jovian planets much larger than the Terrestrial planets?

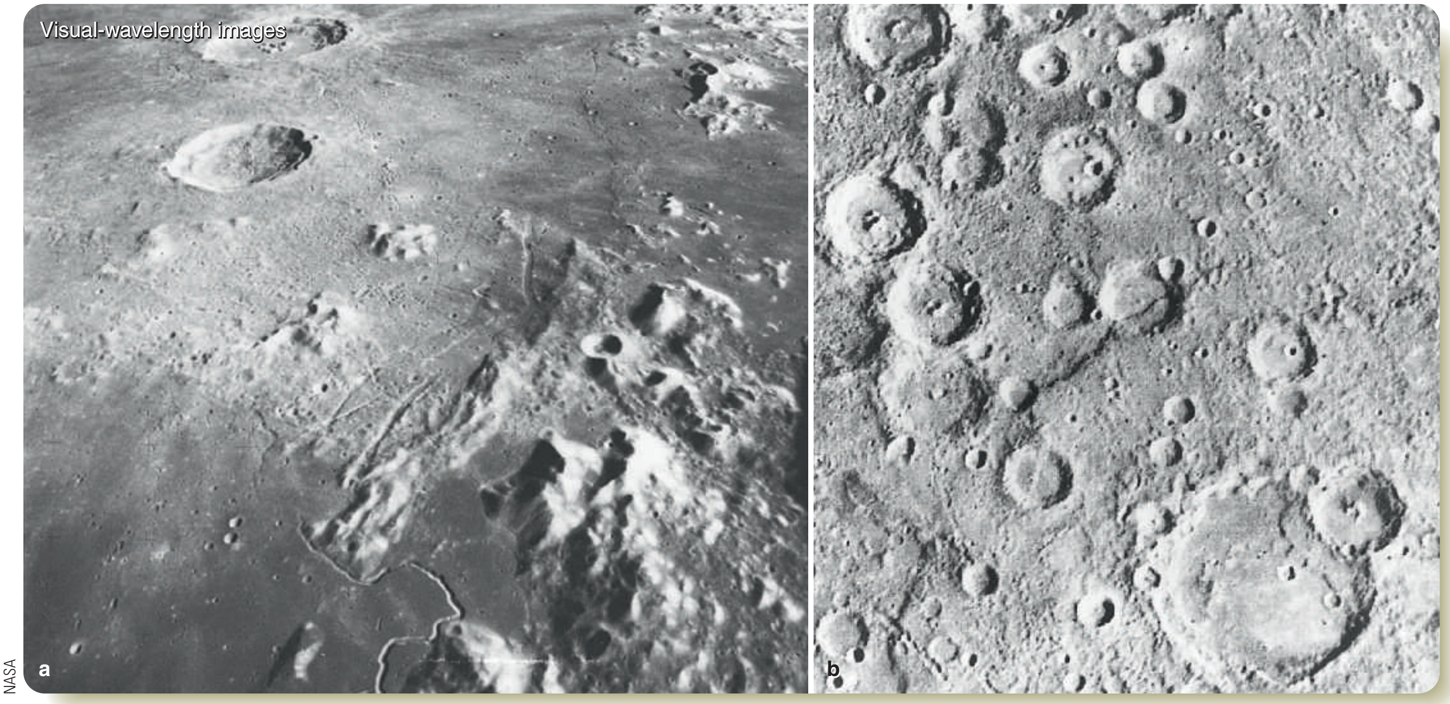
This is an opportunity for you to build an argument that closely analyzes the solar nebula theory. Planets begin forming from solid bits of matter, not from gas. Consequently, the kind of planet that forms at a given distance from the sun depends on the kind of substances that can condense out of the gas there to form solid particles. In the inner parts of the solar nebula, the temperature was so high that most of the gas could not condense to form solids. Only metals and silicates could form solid grains, and the innermost planets grew from this dense material. Much of the mass of the solar nebula consisted of hydrogen, helium, water vapor, and other gases, and they were present in the inner solar nebula but couldn't form solid grains. The small Terrestrial planets could grow only from the solids in their zone, not from the gases, so the Terrestrial planets are small and dense.

In the outer solar nebula, the composition of the gas was the same, but it was cold enough for water vapor and other simple molecules containing hydrogen to condense to form ice grains. Because hydrogen was so abundant, lots of ice could form. The outer planets grew from large amounts of ice combined with small amounts of metals and silicates. Eventually the outer planets grew massive enough that they could begin to capture gas directly from the nebula, and they became the hydrogen- and helium-rich Jovian worlds.

The condensation sequence combined with the solar nebula theory gives you a way to understand the difference between the Terrestrial and Jovian planets. Now expand your argument: **Why do some astronomers argue that the formation of the Jovian planets is a problem that needs further explanation?**

■ **Figure 8-10**

Every old, solid surface in the solar system is scarred by craters. (a) Earth's moon is scarred by craters ranging from basins hundreds of kilometers in diameter down to microscopic pits. (b) The surface of Mercury, photographed by a passing spacecraft, shows vast numbers of overlapping craters.



8-4 Planets Orbiting Other Stars

Are there other planetary systems? The evidence says yes. Do they contain planets like Earth? The evidence so far is incomplete.

Planet-Forming Disks

Both visible- and radio-wavelength observations detect dense disks of gas and dust orbiting young stars. For example, at least 50 percent of the stars in the Orion Nebula are surrounded by such disks (Figure 8-2). A young star is detectable at the center of most disks, and astronomers can measure that the disks contain many Earth masses of material in a region a few times larger in diameter than our solar system. The Orion star-forming region is only a few million years old, so planets may not have formed in these disks yet. Furthermore, the intense radiation from nearby hot stars is evaporating the disks so quickly that planets may never have a chance to grow large. The important point for you to consider is that so many young stars have disks. Evidently, disks of gas and dust are a common feature around stars that are forming.

The *Hubble Space Telescope* also can detect dense disks around young stars in another way. Some disks show up in silhouette against the nebulae that surround the newborn stars (■ Figure 8-11). These disks are related to the formation of bipolar flows in that they focus the gas flowing away from a young star into two jets shooting in opposite directions.

In addition to these dense, hot disks forming planets around young stars, infrared astronomers have found cold, low-density dust disks around stars much older than the newborn stars in Orion, old enough to have finished forming. These tenuous dust disks are sometimes called **debris disks** because they are evidently made of dusty debris produced in collisions among small bodies orbiting those stars that resemble asteroids, comets, and KBOs in our solar system, rather than dust left over from an original protostellar disk. That conclusion is based on calculations showing that the observed dust would be removed by radiation pressure in a much shorter time than the ages of those stars, meaning the dust there now must have been created relatively recently. Our own solar system contains such "second-generation" dust produced by asteroids, comets and Kuiper belt objects. Astronomers consider the solar system's Kuiper belt extending beyond the orbit of Neptune as an example of an old debris disk.

Some examples of debris disks are around the stars Beta Pictoris and Epsilon Eridani (■ Figure 8-12). The dust disk around Beta Pictoris, an A-type star more massive and luminous than the sun, is about 20 times the diameter of our solar system. The dust disk around Epsilon Eridani, which is a K-type star somewhat smaller than the sun, is similar in size to the solar system's Kuiper belt. Like most of the other known low-density disks, both of these examples have central zones with even lower density. Those inner regions are understood to be places where planets have finished forming and have swept up most of the construction material.

Planets orbiting stars with debris disks have not yet been definitely detected, but the presence of dust with short lifetimes around old stars indicates that larger bodies such as asteroids and comets must be present as sources of the dust. If objects of those sizes are there, then it is likely planets are also orbiting those stars.

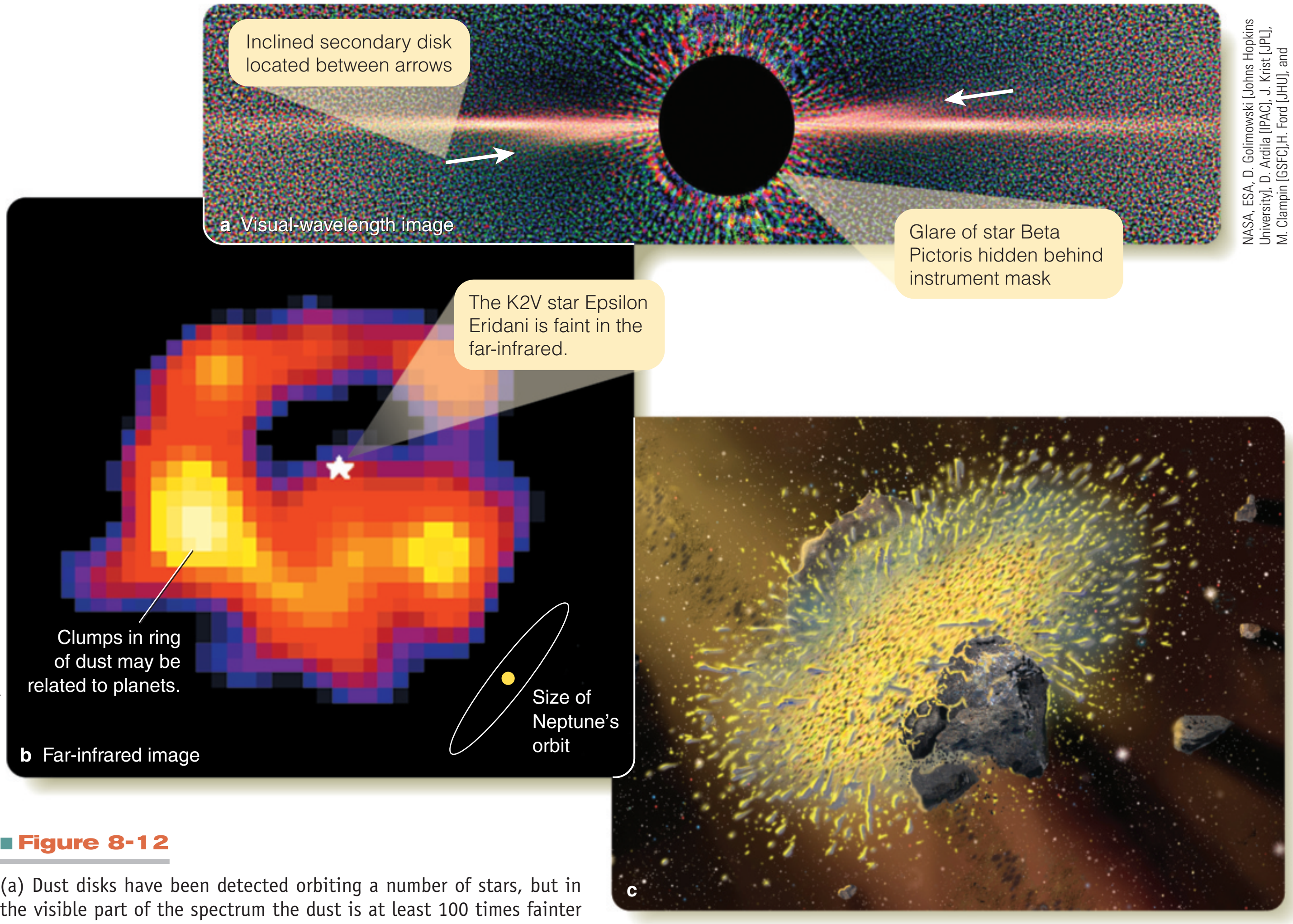
Infrared observations reveal that Favorite Star Vega, easily visible in the Northern Hemisphere summer sky, also has a debris disk, and detailed studies show that most of the dust particles in that disk are tiny. Radiation pressure from Vega should blow away small dust particles quickly, so astronomers conclude that the dust being observed now must have been produced by a big event like the collision of two large planetesimals within the last million years (Figure 8-12c). Fragments from that collision are still smashing into each other now and then and producing more dust, continuing to enhance the debris disk. This effect has also been found in the disk around other mature stars. Such smashups probably happen rarely in a dust disk, but when they happen, they make the disk very easy to detect.

Notice the difference between the two kinds of planet-related disks that astronomers have found. The low-density dust disks such as the ones around Beta Pictoris, Epsilon Eridani, and Vega are produced by dust from collisions between remnant planetesimals. Such disks are evidence that planetary systems have already formed (■ Figure 8-13). In comparison, the dense disks of gas and dust such as those seen round the stars in Orion are sites where planets could be forming right now.

■ **Figure 8-11**

Dark bands (indicated by arrows) are edge-on disks of gas and dust around young stars seen in *Hubble Space Telescope* near-infrared images. Planets may eventually form in these disks. These systems are so young that material is still falling inward and being illuminated by light from the stars.

D. Padgett (IPAC/Caltech), W. Brandner (IPAC), K. Stapelfeldt (JPL) and NASA



NASA, ESA, D. Golimowski [Johns Hopkins University], D. Ardila [IPAC], J. Krist [JPL], M. Clampin [GSFC],H. Ford [JHU], and G. Illingworth [UCO/Lick Obs.]

Joint Astronomy Center

J. Lomberg/Gemini Observatory

© Cengage Learning 2014

Figure 8-12

(a) Dust disks have been detected orbiting a number of stars, but in the visible part of the spectrum the dust is at least 100 times fainter than the stars, which must be hidden behind masks to make the dust detectable. The second faint inclined disk in the Beta Pictoris system may show the orbital plane of a massive planet. (b) At far-infrared wavelengths, the dust in debris disks can be much brighter than the central star. Warps and clumps in these disks suggest the gravitational influence of planets. (c) Collisions between asteroids are rare events, but they generate lots of dust and huge numbers of fragments, as in this artist's conception. Further collisions between fragments can continue to produce dust. Because such dust is blown away quickly, astronomers treat the presence of dust as evidence that objects of at least planetesimal size are also present.

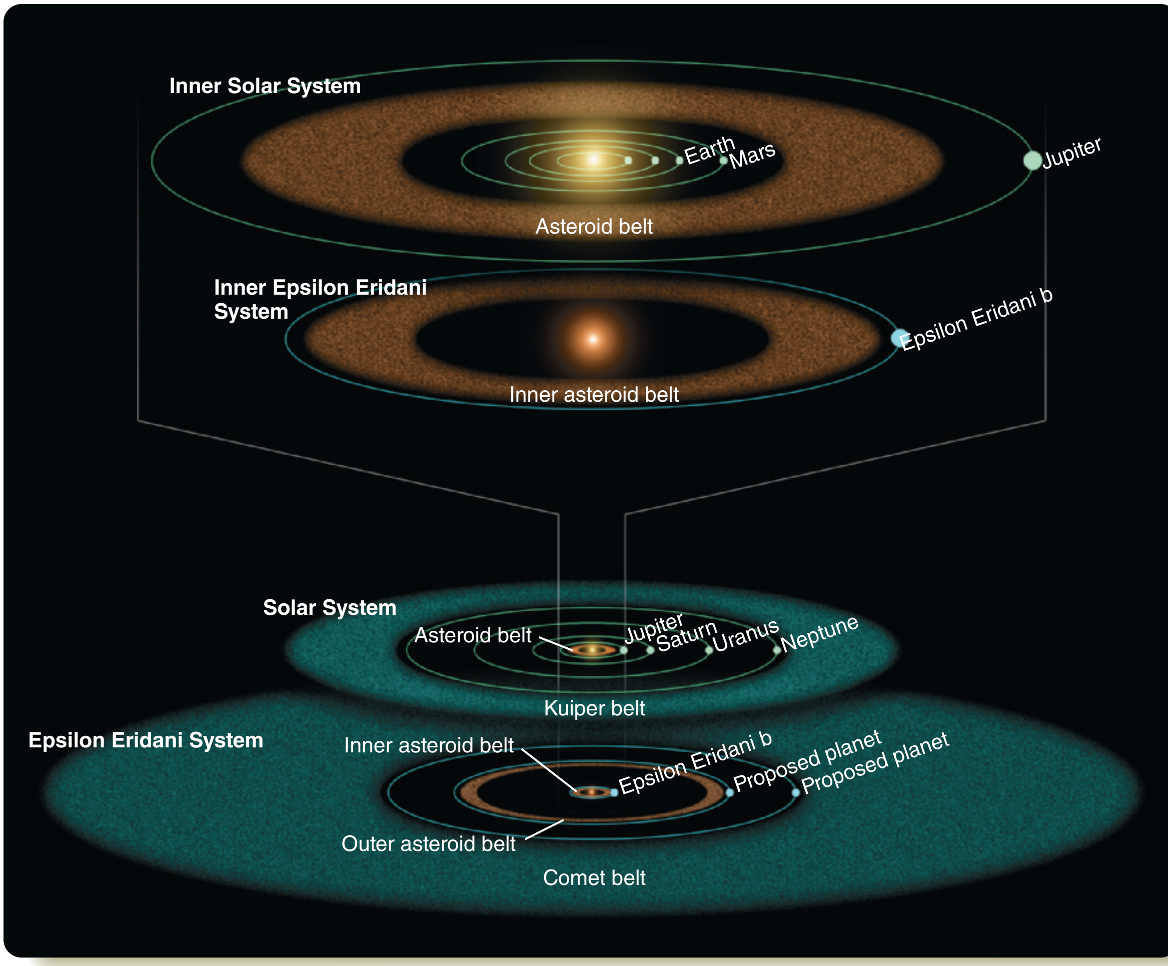
Observing Extrasolar Planets

A planet orbiting another star is called an **extrasolar planet.** Such a planet would be quite faint and difficult to see so close to the glare of its star. But there are ways to find these planets. To understand one important way, all you have to do is imagine walking a dog.

You will remember that Earth and the moon orbit around their common center of mass, and two stars in a binary system orbit around their center of mass. When a planet orbits a star, the star moves very slightly as it orbits the center of mass of the planet–star system. Think of someone walking a poorly trained dog on a leash; the dog runs around pulling on the leash, and even if it were an invisible dog, you could plot its path by watching how its owner was jerked back and forth. Astronomers can detect a planet orbiting another star by watching how the star moves as the planet tugs on it.

The first planet detected this way was discovered in 1995 orbiting the sunlike star 51 Pegasi. As the planet circles the star, the star wobbles slightly, and that very small motion of the star is detectable by Doppler shifts in the star's spectrum (look back to Chapter 6) (■ Figure 8-14a). From the motion of the star and estimates of the star's mass, astronomers deduced that the planet has at least half the mass of Jupiter and orbits only 0.05 AU from the star. Half the mass of Jupiter amounts to 160 Earth masses, so this is a large planet. Note also that it orbits very close to its star, much closer than Mercury orbits around our sun.

Astronomers were not surprised by the announcement that a planet had been found orbiting a sun-like star. For years, astronomers had assumed that many stars had planets. Nevertheless, they carefully tested the data and made further observations that confirmed the discovery. In fact, more than 800 planets have been discovered in this way, including at least



Artist's conception based on *Spitzer Space Telescope* and Caltech Submillimeter Observatory observations by D. Backman et al.; NASA/JPL-Caltech/R. Hurt

Figure 8-13

Dust in debris belts around older main-sequence stars indicates ongoing collisions of remnant planetesimals such as asteroids and comets. Such activity in our solar system is ultimately driven by the gravitational influence of planets. The locations of debris belt edges may be defined by adjacent orbits of planets. The inferred architecture of the Epsilon Eridani planetary system is shown in comparison with our solar system.

seven planets orbiting the star HD 10180, six around Gliese 581, and six around Kepler-11 true planetary systems. Over 100 such multiple-planet systems have been found so far.

Another way to search for planets is to look for changes in the brightness of the star when the orbiting planet crosses in front of the star, called a **transit.** The decrease in light during a planet transit is very small, but it is detectable, and astronomers have used this technique to find many planets. From the amount of light lost, astronomers can tell that these planets with Jovian masses have Jovian diameters, and thus Jovian densities and compositions.

The *Spitzer Space Telescope* has detected infrared radiation from two large hot planets already known from star wobble Doppler shifts. As these planets orbit their parent stars, the amount of infrared radiation from each system varies. When the planets pass behind their parent stars, the total infrared brightness of the systems noticeably decreases. These measurements confirm the existence of the planets and indicate their temperatures and sizes.

Notice how the techniques used to detect these planets resemble techniques used to study binary stars (to be discussed in Chapter 13). Most of the planets were discovered using the same observational methods used to study spectroscopic binary systems, but some were found by observing the systems as eclipsing binaries. In contrast, a few extrasolar planets have been found by a technique called **microlensing,** in which an extrasolar planet passes precisely between Earth and a background star, briefly magnifying the distant star's brightness by gravitational lensing.

The planets discovered so far tend to be massive and have short orbital periods because lower-mass planets or longer-period planets are harder to detect. Low-mass planets don't tug on their stars very much, and present-day spectrographs can't detect the very small velocity changes that these gentle tugs produce. Planets with longer periods are harder to detect because astronomers have not been making high-precision observations for a long enough time. Jupiter takes 11 years to circle the sun once, so it

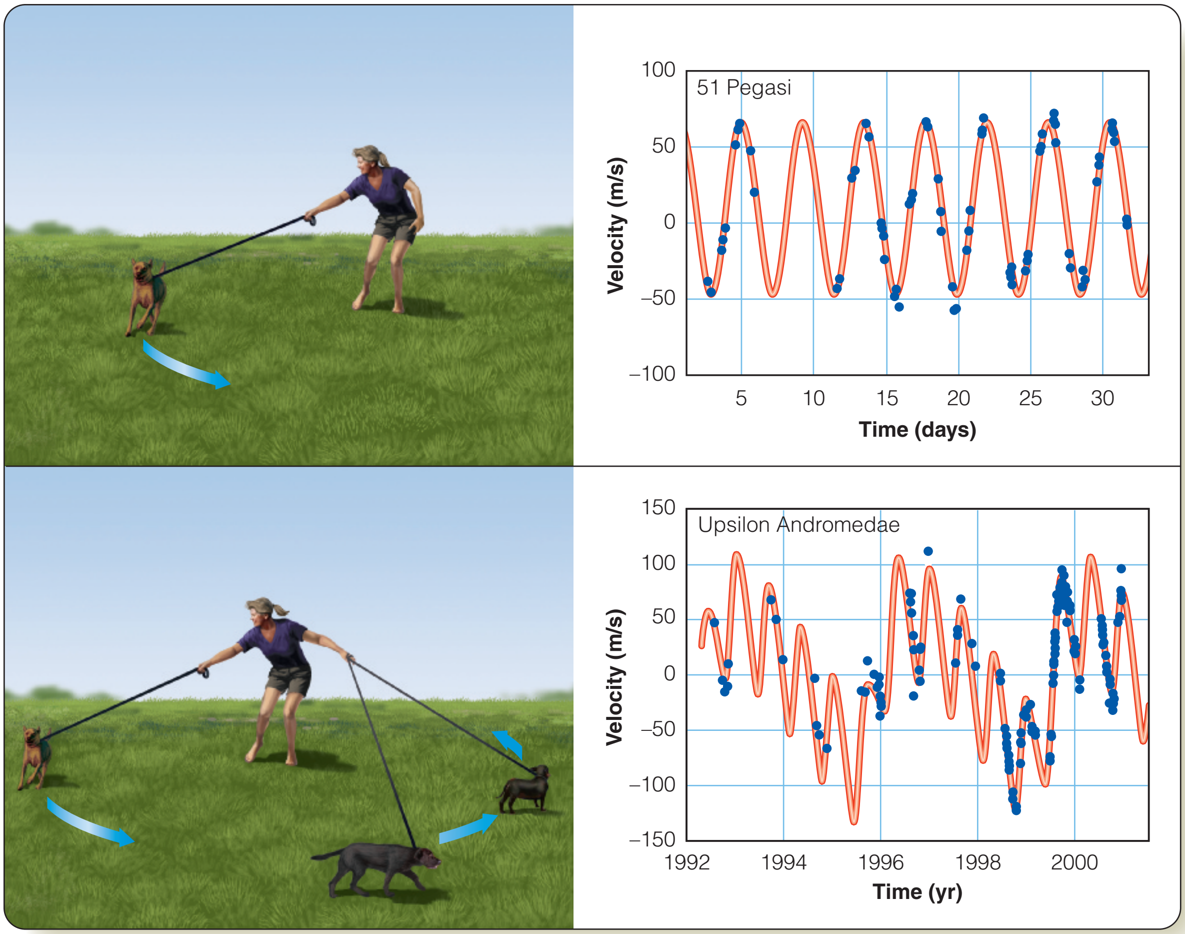


Figure 8-14

Just as someone walking a lively dog is tugged around, the star 51 Pegasi is pulled back and forth by the gravity of the planet that orbits it every 4.2 days. The wobble is detectable in precision observations of the star's Doppler shift. Someone walking three dogs is pulled about in a more complicated pattern, and you can see something similar in the Doppler shifts of Upsilon Andromedae, which is orbited by four planets detected so far. The influence of two of those planets has been removed in this graph to reveal the orbital influences of the other two planets.

will take years for astronomers to see the longer-period wobbles produced by planets lying farther from their stars. You should not be surprised that the first planets discovered are massive and have short orbital periods.

The new planets seem puzzling for several reasons. As you have learned from our solar system, the large planets formed farther from the sun where the solar nebula was colder and ices could condense. How could big planets near their stars (called "**hot Jupiters**") have formed? Theoretical calculations indicate that planets forming in an especially dense disk of matter could spiral inward as they sweep up gas and planetesimals. That means it is possible for a few planets to become the massive, short-period planets that are detected most easily.

Another puzzle is that many of the newly discovered extrasolar planets have elliptical orbits. A simple interpretation of the solar nebula theory would predict that planets generally should have nearly circular orbits as they do in our solar system. Theorists point out, however, that planets in some young planetary systems can interact with each other and can be thrown into elliptical orbits. This effect is probably rare in planetary systems, but astronomers have found these extreme systems more easily because they tend to produce easily detected wobbles.

As astronomers continue to refine their instruments to detect smaller velocity shifts in stars, they find lower-mass planets. A planet with a minimum mass only 1.3 times that of Earth was confirmed in 2012 by careful analysis of Doppler shift data. It orbits its solar-type parent star in 1.2 days at a distance of 0.02 AU.

Actually getting an image of a planet orbiting another star is about as easy as photographing a bug crawling on the bulb of a searchlight miles away. Planets are small and dim and get lost in the glare of the stars they orbit. Nevertheless, during 2008, astronomers managed to image three planets around the A-type

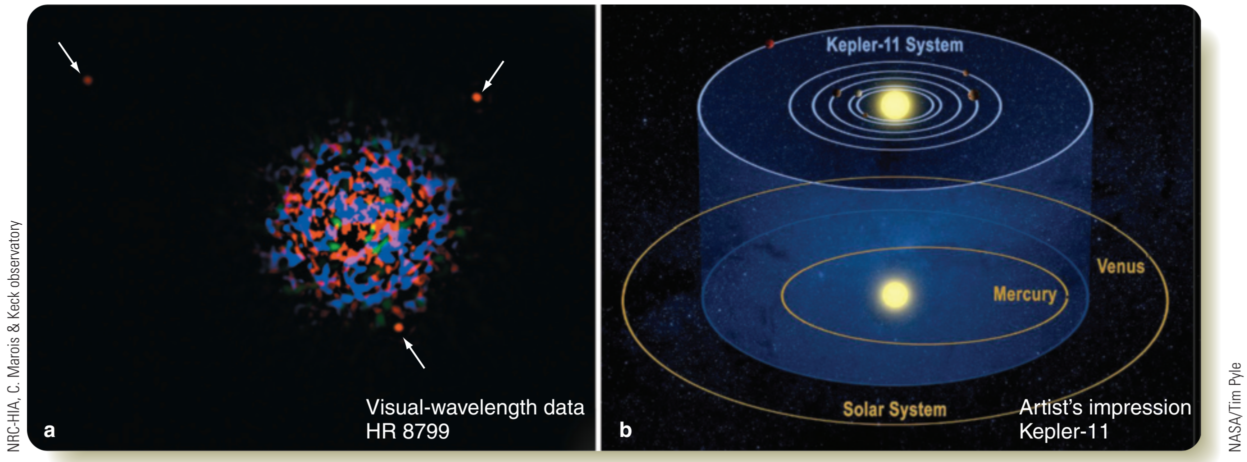


Figure 8-15

Images of three planets at apparent distances of 25, 40, and 70 AU from the star HR 8799. (b) Artist's conception comparing the Kepler-11 planetary system with our solar system. Observations indicate that the six planets so far discovered around Kepler-11 orbit their star in approximately the same plane, like the planets in our solar system.

star HR 8799, using adaptive optics and specially designed cameras on the Gemini and Keck telescopes atop Mauna Kea (Figure 8-15a).

Many research groups around the world are conducting surveys for extrasolar planets. The *Kepler* space telescope, launched in 2009, is "staring" continuously at 150,000 stars searching for extrasolar planets by the transit method. As of this writing, the *Kepler* team has identified more than 1000 candidate planets, some smaller in diameter than Earth. One planetary system discovered by *Kepler* and confirmed by ground-based Doppler shift measurements has six super-Earth planets, more massive than Earth but less massive than Uranus (Figure 8-15b).

Eventually, observatories can be expected to image Jovian and even Terrestrial planets directly around sunlike stars. The discovery of extrasolar planets gives astronomers added confidence in the solar nebula theory. The theory predicts that planets are common, and now astronomers are finding thousands of them.

SCIENTIFIC ARGUMENT

Why are debris disks evidence that planets have already formed?
Sometimes a good scientific argument combines evidence, theory, and an astronomer's past experience, a kind of scientific common sense. Certainly the cold debris disks seen around stars like Vega are not places where planets are forming. They are not hot enough or dense enough to be young disks. Rather, the debris disks must be older, and the dust is being produced by collisions among comets, asteroids, and Kuiper belt objects. Small dust particles would be blown away or destroyed relatively quickly, so these collisions must be a continuing process. The successful solar nebula theory gives astronomers reason to believe that where you find comets, asteroids, and Kuiper belt objects, you should also find planets, so the debris disks seem to be evidence that planets have already formed in such systems.

Now build a new argument. **What direct evidence can you cite that planets orbit other stars?**

What Are We? Planet-Walkers

The matter you are made of came from the big bang, and it has been cooked into a wide variety of atoms inside stars. Now you can see how those atoms came to be part of Earth. Your atoms were in the cloud of gas that formed the solar system 4.6 billion years ago, and nearly all of that matter contracted to form the sun, but a small amount left behind in a disk formed planets. In the process, your atoms became part of Earth.

You are a planet-walker, and you have evolved to live on the surface of Earth. Are there other beings like you in the universe? Now you know that planets are common, and you can reasonably suppose that there are more planets in the universe than there are stars. However complicated the formation of the solar system was, it is a common process, so there may indeed be more planet-walkers living on other worlds.

But what are those distant planets like? Before you can go very far in your search for life beyond Earth, you need to explore the range of planetary types. It is time to pack your spacesuit and voyage out among the planets of our solar system, visit them one by one, and search for the natural principles that relate planets to each other. Although the formation of the solar system was a complicated process, there may in fact be more planet-walkers living on other worlds.

Study and Review

Summary

- Hypotheses for the origin of the solar system have been either catastrophic or evolutionary. **Catastrophic hypotheses (p. 142)** depend on a rare event such as the collision of the sun with another star. **Evolutionary hypotheses (p. 142)** propose that the planets formed by gradual, natural processes. The evidence now strongly favors the **solar nebula theory (p. 142)**, an evolutionary scenario.
- Modern astronomy reveals that all the matter in the universe, including our solar system, was originally formed as hydrogen and helium in the big bang. Atoms heavier than helium were cooked up in nuclear reactions in later generations of stars. The sun and planets evidently formed from a cloud of gas and dust in the interstellar medium.
- The solar nebula theory proposes that the planets formed in a disk of gas and dust around the protostar that became the sun. Hot disks of gas and dust have been detected around many protostars and are believed to be the kind of disk in which planets could form.
- The solar system is disk shaped, with all the planets orbiting nearly in the same plane. The orbital revolution of all the planets, the rotation of most of the planets on their axes, and the revolution of most of their moons are all in the same direction, counterclockwise as seen from the north.
- The planets are divided into two groups. The inner four planets are **Terrestrial planets (p. 146)**—small, rocky, dense, Earth-like worlds. The next four outward are Jupiter-like **Jovian planets (p. 146)** that are large and low density.
- All four of the Jovian worlds have ring systems and large families of moons. The Terrestrial planets have no rings and few moons.
- Most of the **asteroids (p. 145)**, small, irregular, rocky bodies, are located between the orbits of Mars and Jupiter.
- The **Kuiper belt (p. 148)** is composed of small, icy bodies called **Kuiper belt objects (KBOs) (p. 148)** that orbit the sun beyond the orbit of Neptune.
- **Comets (p. 148)** are icy bodies that pass through the inner solar system along long elliptical orbits. As the ices vaporize and release dust, effects of the solar wind plus **radiation pressure (p. 148)** cause the comet to develop a tail that points approximately away from the sun.
- **Meteoroids (p. 149)** that fall into Earth's atmosphere are vaporized by friction and are visible as streaks of light called **meteors (p. 148)**. Larger and stronger meteoroids may survive to reach the ground, where they are called **meteorites (p. 149)**.
- The age of a rocky body can be found by radioactive dating, based on the decay **half-life (p. 149)** of radioactive atoms. The oldest rocks from Earth, the moon, and Mars have ages over 4 billion years. The oldest objects in our solar system are some meteorites that have ages of 4.6 billion years. This is taken to be the age of the solar system.
- **Condensation (p. 152)** in the solar nebula converted some of the gas into solid bits of matter. According to the **condensation sequence (p. 152)**, the inner part of the solar nebula was so hot that only metals and rocky materials could form solid grains. The dense Terrestrial planets grew from those solid particles and did not include many ices or low-density gases. Comparing the **uncompressed densities (p. 151)** of the Terrestrial planets shows that the inner Terrestrial planets have higher densities than the outer ones because of the condensation sequence.
- The outer solar nebula, beyond the **ice line (p. 152)**, was cold enough for large amounts of ices as well as metals and rocky minerals to form solid particles. The Jovian planets grew rapidly and incorporated large amounts of low-density ices and gases. Further evidence that the condensation sequence was important in the solar nebula is found in the high densities of the Terrestrial planets relative to the Jovian planets.
- The solar nebula theory predicts that as solid particles became larger they would no longer grow efficiently by condensation but could continue growing by **accretion (p. 152)** to form billions of **planetesimals (p. 152)**.
- Planets begin growing by accretion of solid material into **protoplanets (p. 153)**. Once a protoplanet approaches about 15 Earth masses, it can begin growing by **gravitational collapse (p. 153)** as it pulls in gas from the solar nebula.
- The Terrestrial planets may have formed slowly from the accretion of planetesimals of similar composition. Radioactive decay plus **heat of formation (p. 153)** then melted each planet's interior to cause **differentiation (p. 153)** into layers of differing density. In that scenario, Earth's early atmosphere was probably supplied by a combination of **outgassing (p. 154)** from Earth's interior and planetesimal impacts.
- Disks of gas and dust around protostars may not last long enough to form Jovian planets by accretion and then by gravitational collapse. This is referred to as the **Jovian problem (p. 155)**. Some models suggest the Jovian planets could have formed more rapidly by **direct collapse (p. 155)**, skipping the condensation and accretion steps.
- In addition to evaporation by the light from hot nearby stars and the gravitational influence of passing stars, the solar nebula was eventually cleared away by radiation pressure, the solar wind, and the sweeping up or ejection of debris by the planets.
- All of the old surfaces in the solar system were heavily cratered during the **heavy bombardment (p. 157)** by debris that filled the solar system when it was young, left over from the formation of the planets.
- Cold dust disks known as **debris disks (p. 158)** around main-sequence stars appear to be produced by dust released by collisions among comets, asteroids, and Kuiper belt objects. Such disks may be signs that planets have already formed in those systems.
- Planets orbiting other stars, called **extrasolar planets (p. 159)**, have been detected by the way they tug their stars about, creating small Doppler shifts in the stars' spectra. Planets have also been detected in **transits (p. 160)** as they cross in front of their star and partly block the star's light. A few planets have been detected when they orbited behind their star and their infrared radiation was cut off. A few more have been detected by gravitational **microlensing (p. 160)**.
- Nearly all extrasolar planets found so far are massive, Jovian worlds orbiting close to their parent stars, called **hot Jupiters (p. 161)**. Lower-mass Terrestrial planets are harder to detect but may be common. Astronomers speculate that hot Jupiters formed beyond the ice lines of their respective planetary systems and then spiraled inward to their current locations as they swept up protoplanetary disk material.

Review Questions

1. What produced the helium now present in the sun's atmosphere? In Jupiter's atmosphere? In the sun's core?
2. What produced the iron and heavier elements like gold and silver in Earth's core and crust?
3. What evidence can you cite that disks of gas and dust are common around young stars?
4. According to the solar nebula theory, why is the Earth's orbit nearly in the plane of the sun's equator?

5. Why does the solar nebula theory predict that planetary systems are common?
6. What is the evidence that the solar system formed about 4.6 billion years ago?
7. If you visited another planetary system, would you be surprised to find planets older than Earth? Why or why not?
8. Why is almost every solid surface in our solar system scarred by craters?
9. What is the difference between condensation and accretion?
10. Why don't Terrestrial planets have rings like the Jovian planets?
11. How does the solar nebula theory help you understand the location of asteroids?
12. How does the solar nebula theory explain the dramatic density difference between the Terrestrial and Jovian planets?
13. What does the term *differentiated* mean when applied to a planet? Would you expect to find that planets are usually differentiated? Why?
14. What processes cleared the nebula away and ended planet building?
15. Why would astronomically short lifetime of gas and dust disks around protostars pose a problem in understanding how the Jovian planets formed? What modification of the solar nebula theory might solve this problem?
16. What is the difference between the dense hot disks seen around some stars and the low-density cold disks seen around some other stars?
17. What evidence can you cite that planets orbit other stars?
18. Why is the existence of "hot Jupiters" puzzling? What is the current hypothesis for how they formed?
19. **How Do We Know?** The evidence is overwhelming in support of the idea that the Grand Canyon was dug over a span of millions of years by the erosive power of the Colorado River and its tributary streams. Is that a catastrophic theory or an evolutionary theory?
20. **How Do We Know?** How can scientists know anything about how the solar system formed, given that there was nobody there to witness those events?

Discussion Questions

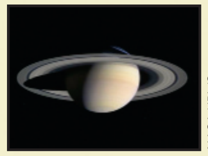
1. If you visited some other planetary system while the planets were forming, would you expect to see the condensation sequence at work, or do you think that was most likely unique to our solar system? How do the properties of the extrasolar planets discovered so far affect your answer?
2. In your opinion, do most planetary systems have asteroid belts? Would all planetary systems show evidence of an age of heavy bombardment?
3. If the solar nebula hypothesis is correct, then there are probably more planets in the universe than stars. Do you agree? Why or why not?

Problems

1. If you observed the solar system from the nearest star (distance = 1.3 parsecs), what would the maximum angular separation be between Earth and the sun? (*Note:* 1 pc is 2.1×10^5 AU.) (*Hint:* Remember to use the small-angle formula, Chapter 3.)
2. The brightest planet in our sky is Venus, which is sometimes as bright as apparent magnitude −4 when it is at a distance of about 1 AU. How many times fainter would it look from a distance of 1 parsec? What would its apparent magnitude be? Note that 1 pc is 2.06×10^5 AU. (*Hints:* Look ahead to the inverse square law in Chapter 13; also, review the definition of magnitudes in Chapter 2.)
3. What is the smallest-diameter crater you can identify in the photo of Mercury on page 146? (*Hint:* See Appendix Table A-10, "Properties of the Planets," to find the radius of Mercury in kilometers.)
4. A sample of a meteorite has been analyzed, and the result shows that out of every 1000 nuclei of potassium-40 originally in the meteorite, only 200 have not decayed. How old is the meteorite? (*Hint:* See Figure 8-6.)
5. In Table 8-2, which object's observed density differs least from its uncompressed density? Why?
6. What composition might you expect for a planet that formed in a region of the solar nebula where the temperature was about 100 K?
7. Suppose that Earth grew to its present size in 1 million years through the accretion of particles averaging 100 grams each. On the average, how many particles did Earth capture per second? (*Hint:* See Appendix Table A-10 to find Earth's mass.)
8. If you stood on Earth during its formation, as described in Problem 7, and watched a region covering 100 m^2, how many impacts would you expect to see in an hour? Note that the surface area of a sphere is $4\pi r^2$. (*Hint:* Assume that Earth had its present radius, given in Appendix Table A-10.)
9. The velocity of the solar wind is roughly 400 km/s. How long does it take to travel from the sun to Earth?

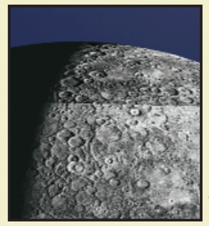
Learning to Look

1. What do you see in the image below that indicates this planet formed far from the sun?



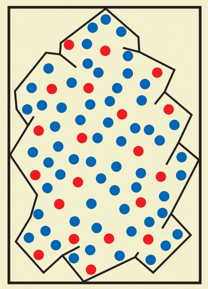
NASA/JPL/Space Science Institute

2. Why do astronomers conclude that the surface of Mercury, shown below, is old? When did the majority of those craters form?



NASA

3. In the mineral specimen represented below, radioactive atoms (red) have decayed to form daughter atoms (blue). How old is this specimen in half-lives? (See Figure 8-6).



© Cengage Learning 2014

Great Debates

1. *Are We Recycled Humans?* The human race is much younger than the age of the solar system (refer back to Section 1-2). Some objects on Earth, such as the Egyptian pyramids, seemed to have been built by civilizations more intelligent than the average human intelligence at that time. Could the human race have been driven to extinction several times during bombardment periods during the solar system's lifetime? If so, could you be a recycled human?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers in your debate.
   b. What's the evidence? Find additional information that supports your claim, such as places on Earth that indicate past advanced civilizations. Embed and discuss pictures of such places on Earth.
   c. Cite your sources.

2. *Are You an Ostrich or Bear?* Suppose many governments of the world have charted a large asteroid's path and know that next year the asteroid is likely to hit Earth. Not enough time is available to alter the asteroid's path or otherwise avoid an impact. Should the governments inform the public of the impact? If so, how much advance notice should governments give? Should governments inform the public of human survival chances? Would you want to know? People, in general, fall into two camps—the ostrich and the bear. The ostrich hides its head in the sand during times of stress, whereas the bear charges the danger. Which one are you, the ostrich or the bear?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers in your debate.
   b. What's the evidence? Find additional information about asteroid impacts that supports your claim.
   c. Cite your sources.

3. *Three Kinds of Planets?* In astronomy textbooks, usually two kinds of planets are discussed, Jovian and Terrestrial. But some argue that the Jovian planets can be subdivided into gas/liquid giants and ice giants, which have less hydrogen and helium. Should astronomy students at any school level learn about three types of planets or two types of planets in the formation of the solar system?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers in your debate.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

4. *Oort Cloud Comets*. Long-period comets are known to exist, but are they from the Oort cloud? If a new model has gained the confidence of the scientific community to confirm Oort cloud comets, should that person or persons who developed the model receive the Nobel Prize in physics for the discovery? How would you vote?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers in your debate.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

5. *Another Armageddon?* If a comet or asteroid impact with Earth is likely, should secondary students by informed by teachers and principles, or should the students be sent home without a reason, leaving parents to explain the Earth's impending doom? Should news like this have a parental guidance rating like the ratings for movies? Should governments set a minimum age on information like this that might be upsetting to children?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers in your debate.
   b. What's the evidence? Find additional information about asteroid impacts that supports your claim.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- The Solar Nebula Theory
- Detection of Exoplanets from Radial Velocity

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 8: Extrasolar Planets

As soon as Earth was understood to be just one more planet orbiting the sun among its siblings, speculation began about the possibility of life on other worlds. When the Copernican model of the solar system with the sun at the center superseded the ancient model of a unique and unmoving Earth at the center of the universe, imaginations were set free to envision beings like us living on every little light in the sky. Some writers went so far as to portray beings living on the sun and other stars. Application of the Copernican principle to its logical conclusion would seemingly imply that nothing whatsoever is central or unique about Earth and us.

But you are learning to prefer information to speculation and argumentation. After all, the Copernican principle does not prove anything. This principle just means scientists prefer hypotheses that don't imply Earth is unique in location or other characteristics. Earth in fact could be unique in some important ways. There is a big difference between assuming Earth is ordinary and endeavoring to find out whether that is true. One important part of the question of Earth's uniqueness is whether our solar system, a set of planets orbiting an ordinary star, is a common or rare phenomenon in the universe.

As you learned in earlier chapters, astronomers found that binary stars orbiting each other follow Newton's laws of motion and gravity just as the planets in our solar system do. Our understanding of stars is based substantially on a combination of observing binary stars and applying the known laws of physics. Many binary star systems are *spectroscopic binaries*, in which we observe only one point of light. But, spectra of that light reveal varying Doppler shifts indicating there are actually two objects orbiting each other. A few binary systems are *eclipsing binaries*, in which the orbit plane is aligned with our line of sight and the stars eclipse each other. You can learn a lot about the stars in an eclipsing binary system by carefully observing the details of the eclipses.

Planets orbiting stars other than our sun are termed "extrasolar planets." As you learned in this chapter, the techniques first used to study binary stars have been refined and made more sensitive to be able to detect extrasolar planets. Extrasolar planetary systems are observationally like lopsided binary star systems with one massive component (a star) and one or more relatively tiny components (planets). You might note that, in our solar system, the sun has 1000 times the mass of the largest planet, Jupiter, so Jupiter is tiny compared with the sun. Nevertheless, Jupiter is more massive than the rest of the planets combined. To an alien astronomer observing from another planetary system, our system might appear to consist of nothing but the sun plus Jupiter.

Section 1 of Virtual Astronomy Lab 8, "Extrasolar Planets," guides you through analyzing the characteristics of some real extrasolar planets discovered via the "Doppler wobble" (spectroscopic binary) technique. Section 2 explores the detection of extrasolar planets by observation of transits, an extension of the eclipsing binary star technique. Section 3 confronts you with an apparent violation of the Copernican principle: The first few hundred extrasolar planetary systems discovered were architecturally unlike our solar system; they have Jovian-mass planets orbiting very close to their parent stars, in some cases much closer than Mercury orbits our sun. Does this mean Earth-like planets are actually rare? Stay tuned. . . Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Through comparative planetology, the moon reveals a great deal about our own beautiful Earth.

NASA/JSC

© Aleksandar Todorovic/Shutterstock.com

9

Earth and Moon: Bases for Comparative Planetology

Guidepost

In the preceding chapter, you learned how our solar system formed as a by-product of the formation of the sun. You also saw how distance from the sun determined the general composition of each planet. In this chapter, you begin your study of individual planets by comparing and contrasting Earth with its moon. On the way, you will answer four important questions:

- **How do the Earth and its moon compare with the other Terrestrial worlds?**
- **How has Earth changed and evolved since it formed?**
- **Why is the moon airless, heavily cratered, and geologically inactive?**
- **How are the histories of Earth and the moon connected?**

Like a foundation constructed before starting a skyscraper, this chapter will establish your basis for studying other worlds by first informing you about Earth and its moon. In the following two chapters you will visit objects that are un-Earthly and un-moonlike, yet, in other ways, share some common features.

We pray for one last landing
On the globe that gave us birth;
Let us rest our eyes on the fleecy skies
And the cool, green hills of Earth.

FROM THE FANCIFUL "NATIONAL ANTHEM" OF EARTH, IN THE NOVELS AND SHORT STORIES OF ROBERT HEINLEIN

Beautiful, beautiful. Magnificent desolation . . .

EDWIN ALDRIN, ON THE MOON

PLANETS ARE MORE ALIKE than they are different; they are described by the same basic principles. You can compare and contrast them to identify those principles and understand planets better. This approach is called **comparative planetology.** You can learn much more that way than by studying the planets separately.

Earth and its moon are ideal starting points for your study because they are respectively the largest and the smallest of the five Terrestrial worlds. Earth is a complex planet with a molten interior that generates a magnetic field. Its crust is active, with moving continents, earthquakes, volcanoes, and ongoing mountain building. Earth's oxygen-rich atmosphere and water-covered surface are unique in the solar system. In contrast, the moon is geologically inactive and has no atmosphere. Nevertheless, the histories of Earth and the moon are intertwined.

You will meet the other three Terrestrial worlds—Mercury, Venus, and Mars—in the next chapter and find that many of their properties fit between the benchmarks set by Earth and the moon.

9-1 A Travel Guide to the Terrestrial Planets

IF YOU VISIT THE CITY OF GRANADA in Spain, you will probably consult a travel guide. If it is a good guide, it will do more than tell you where to find museums and restrooms. It will give you a preview of what to expect. You are beginning a journey to the Earth-like worlds, so you should consult a travel guide and see what is in store.

Five Worlds

In this chapter and the next you are about to visit Earth, Earth's moon, Mercury, Venus, and Mars. It may surprise you that the moon is included in your itinerary. It is, after all, just a natural satellite orbiting Earth and isn't one of the planets. But the moon is a fascinating world of its own, it makes a striking comparison with the other worlds on your list, and its history gives you important information about the history of Earth and the other planets.

■ Figure 9-1 compares the five worlds. The first feature you might notice is diameter. The moon is small, and Mercury is not much bigger. Earth and Venus are large and quite similar in size, but Mars is a medium-sized world. You will discover that size is a critical factor in determining a world's personality. Small worlds tend to be geologically inactive, while larger worlds tend to be active.

Core, Mantle, and Crust

The Terrestrial worlds are made up of rock and metal. Planetary scientists have evidence that they are all differentiated, which means each world is separated into layers of different density, with a dense metallic core surrounded by a less-dense rocky **mantle,** and a low-density crust on the outside.

As you learned in Chapter 8, when the planets formed, their surfaces were subjected to heavy bombardment by leftover planetesimals and debris in the young solar system. You will see lots of craters on these worlds, especially on Mercury and the moon, many of them dating back to the heavy bombardment era. Notice that cratered surfaces are old. For example, if a lava flow covered up some cratered landscape after the end of the heavy bombardment, few craters could be formed later on that surface because most of the debris in the solar system was gone. When you see a smooth plain on a planet, you can be sure that surface is younger than the heavily cratered areas.

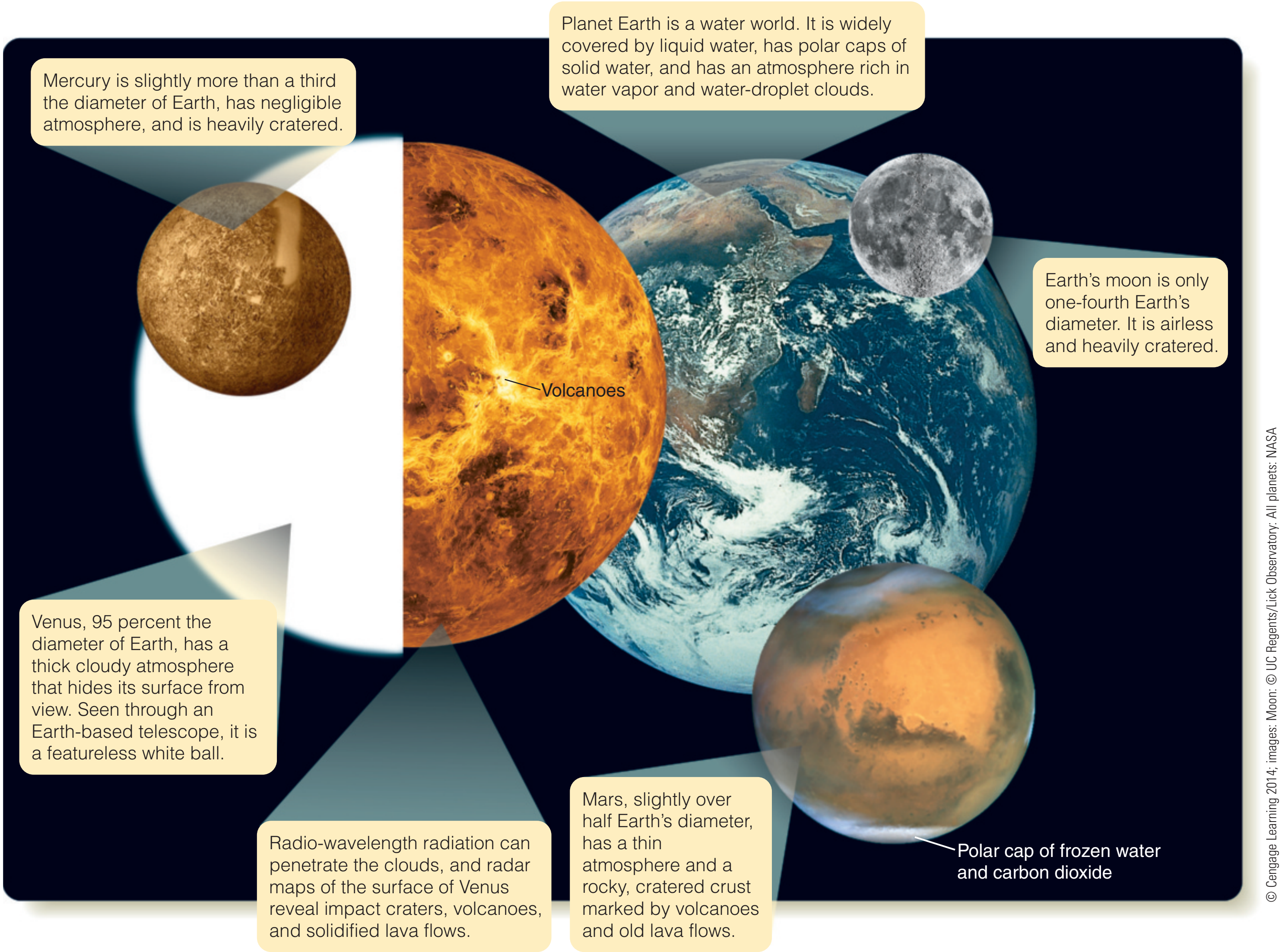
Another important way you can study a planet is by following the energy flow. In the preceding chapter you learned that the heat in the interior of a planet may be partly from radioactive decay and partly left over from the planet's formation, but in any case it must flow outward toward the cooler surface where it is radiated into space. In the process of flowing outward, the heat can cause convection currents in the liquid interior, magnetic fields, plate motions, quakes, faults, volcanism, mountain building, and more. Heat flowing outward through the cooler crust makes a large world like Earth geologically active (**How Do We Know? 9-1**). In contrast, the moon and Mercury, both small worlds, cooled quickly inside, so they have little heat flowing outward now and are relatively inactive.

Atmospheres

When you look at Mercury and the moon in Figure 9-1, you can see their craters, plains, and mountains clearly; they each have no substantial atmosphere to obscure your view. In

■ **Figure 9-1**

Planets in comparison. Earth and Venus are similar in size, but their atmospheres and surfaces are very different. The moon and Mercury are much smaller, and Mars is intermediate in size.



comparison, the surface of Venus is completely hidden by a cloudy atmosphere even thicker than Earth's. Mars, the medium-sized planet, has a relatively thin atmosphere.

You might ponder two questions. First, why do some worlds have atmospheres while some do not? You will discover that both size and temperature are important. The second question is more complex. Where did those atmospheres come from? To answer that question in later chapters, you will have to understand the histories of these worlds.

Four Stages of Planetary Development

Like the other Terrestrial planets, Earth formed from the inner solar nebula about 4.6 billion years ago. Even as it took form, it began to change. There is evidence that Earth and the other Terrestrial planets, plus Earth's moon, passed through four developmental stages (■ Figure 9-2).

The first stage of planetary evolution is *differentiation,* the separation of material into layers according to density. In the

How Do We Know? 9-1

Understanding Planets: Follow the Energy

What causes change? One of the best ways to think about a scientific problem is to follow the energy. According to the principle of cause and effect, every effect must have a cause, and every cause must involve energy. Energy moves from regions of high concentration to regions of low concentration and, in doing so, produces changes. For example, coal burns to make steam in a power plant, and the steam passes through a turbine and then escapes into the air. In flowing from the burning coal to the atmosphere, the heat spins the turbine and makes electricity.

Scientists commonly use energy as a key to understanding nature. A biologist might ask where certain birds get the energy to fly thousands of miles, and a geologist might ask where the energy comes from to power a volcano. Energy is everywhere, and when it moves, whether it is in birds or molten magma, it causes change. Energy is the "cause" in "cause and effect."

You may have learned that the flow of energy from the inside of a star to its surface helps you know how the sun and other stars work. The outward flow of energy supports the star against its own weight, drives convection currents that produce magnetic fields, and causes surface activity such as spots, prominences, and flares. You can understand stars by knowing how energy flows from their interiors to their surfaces and into space.

You can also think of a planet by following the energy. The heat in the interior of a planet may be left over from the formation of the planet, or it may be heat generated by radioactive decay, but it must flow outward toward the cooler surface, where it is radiated into space. In flowing outward, the heat can cause convection currents in the mantle, magnetic fields, plate motions, quakes, faults, volcanism, mountain building, and more.

When you think about any world, be it a small asteroid or a giant planet, think of its interior as a source of heat that flows through the planet's surface into space. If you can follow that energy flow, you can understand a great deal about the world. A planetary astronomer once said, "The most interesting thing about any planet is how its heat gets out."

M. Seeds

Heat flowing out of Earth's interior generates geological activity such as that at Yellowstone National Park.

next section you will learn about evidence that Earth is differentiated, with a dense core plus a less dense mantle and crust. That differentiation is understood to have occurred due to melting of Earth's interior caused by heat from a combination of radioactive decay plus energy released by infalling matter during the planet's formation. Once the interior of Earth melted, the densest materials were able to sink to the core.

The second stage, *cratering,* could not begin until a solid surface formed. The heavy bombardment of the early solar system made craters on Earth just as it did on the moon and other planets. As the debris in the young solar system cleared away, the rate of cratering impacts decreased toward its present low rate.

The third stage, *flooding,* began as radioactive decay continued to heat Earth's interior and caused rock to melt in the upper mantle, where the pressure was lower than in the deep interior. Some of that molten rock welled up through cracks in the crust and flooded the deeper impact basins. Later, as the environment cooled, water fell as rain and flooded the basins to form the first oceans. Note that on Earth, basin flooding was first by lava and later by water.

The fourth stage, *slow surface evolution,* has continued for at least the past 3.5 billion years. Earth's surface is constantly changing as sections of crust slide over and against each other, push up mountains, and shift continents. In addition, moving air and water erode the surface and wear away geological features. Almost all traces of the first billion years of Earth's history have been destroyed by the active crust and erosion.

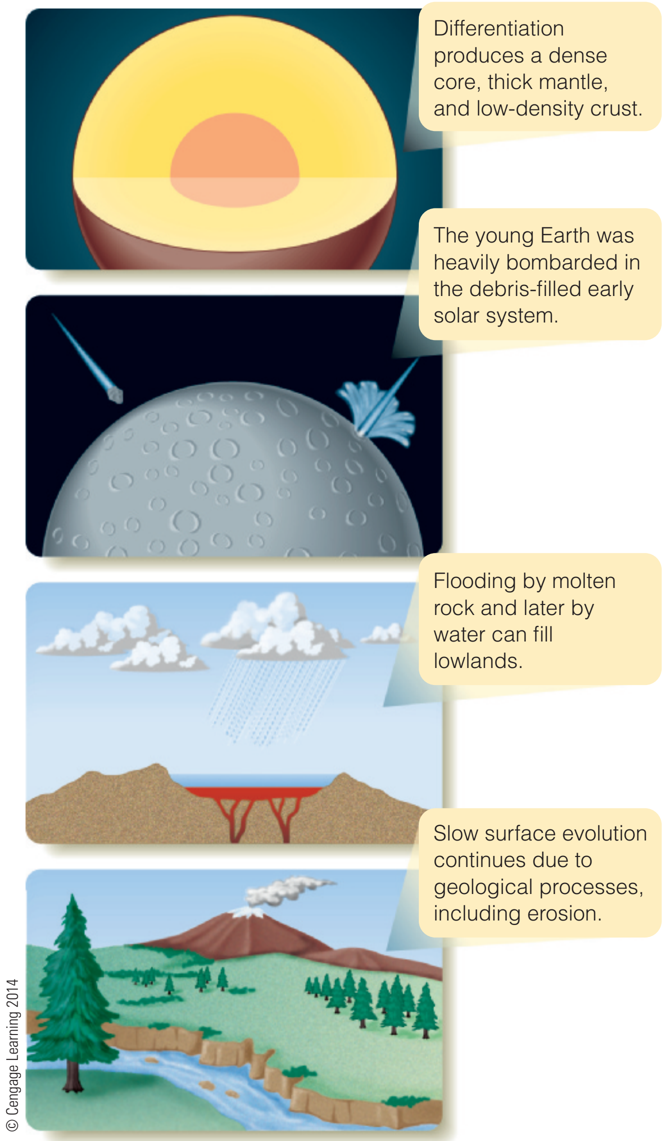
Terrestrial planets pass through these four stages, but differences in mass, temperature, and composition among the planets can emphasize some of those stages relative to the other stages and produce surprisingly different worlds.

SCIENTIFIC ARGUMENT

Why do you expect the inner planets to be high-density worlds? In Chapter 8, you saw how the inner planets formed from hot inner parts of the solar nebula. No ice solidified there, so the inner planets could grow only from particles of rock and metal able to condense from hot gas. Therefore, you expect the inner planets to be made mostly of rock and metal, which are dense materials.

As you visit the Terrestrial planets, you will find craters almost everywhere. **What made all of those craters?**

Four Stages of Planetary Development



■ **Figure 9-2**

The four stages of Terrestrial planet development are illustrated for Earth.

9-2 Planet Earth

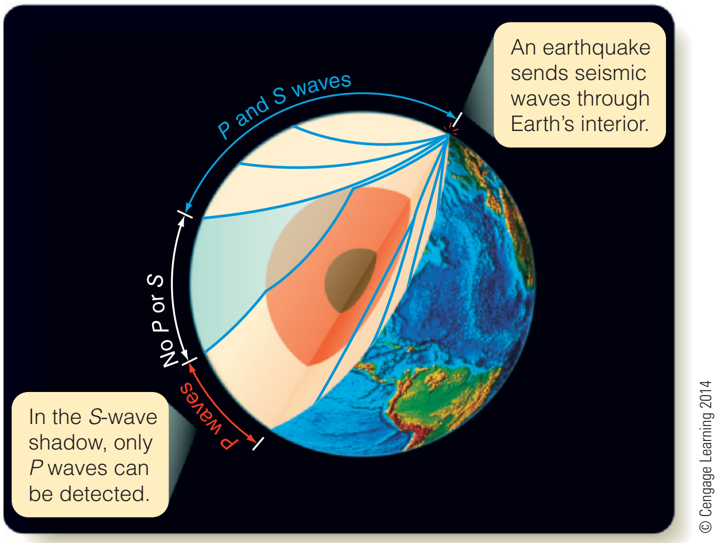
EARTH IS A GOOD STANDARD for comparative planetology (**Celestial Profile 2**, p. 176). Every major process on any rocky world in our solar system is represented in some form on Earth. Nevertheless, Earth is unusual among the planets in our solar system in several ways. More than 70 percent of Earth's surface is covered by liquid water. No other planet in our solar system has liquid water on its surface, although, as you will learn in later chapters, Mars had surface water long ago, Venus probably did, and some moons in the outer solar system show evidence of having liquid water under their surfaces.

Your home planet is special in a second way. Some of the matter on the surface of this world is alive, and a small part of that living matter, including you, is aware. As you will learn in Chapter 20, life currently seems to be absent from the other words in the solar system. No one is sure about all the ways that the presence of living matter has affected the evolution of Earth, but as you will learn later in this chapter the thinking part of the life on Earth, humankind, is actively altering our planet.

Earth's Interior

You have read that there is evidence Earth is differentiated into a dense metallic core with a lower-density rocky mantle and crust. What is the evidence? A very simple argument is based on the fact that Earth's average density, which can be calculated easily from its mass and volume, is 5.5 g/cm^3, but the silicate rocks on Earth's surface have only about half that density. Therefore, a large part of Earth's interior must be made of material much denser than crustal rock. Can more detailed information be determined about Earth's interior?

Although even the deepest mines and wells extend only a few kilometers down and don't even reach the bottom of the crust, exploration of Earth's interior is possible because earthquakes produce vibrations called **seismic waves** that travel through the crust and interior and eventually register on sensitive detectors called **seismographs** all over the world (■ Figure 9-3). Two kinds of seismic waves are important in this exploration. **Pressure *(P)* waves**



■ **Figure 9-3**

P and *S* waves give you clues to the structure of Earth's interior. No direct *S* waves from an earthquake reach the side of Earth opposite their source, indicating that Earth's core is liquid. The size of the *S* wave "shadow" tells you the size of the liquid outer part of the core.

travel as a series of compressions and uncompressions, like sound, and can pass through a liquid. In contrast, the **shear *(S)* waves** travel as a type of side-to-side vibration that can't pass through a liquid. When an earthquake occurs, no direct *S* waves pass through the core to register on seismographs on the opposite side of Earth, as if the core were casting a shadow (Figure 9-3). The absence of *S* waves shows that the core is mostly liquid, and the size of the *S*-wave shadow shows that the outer boundary of the core lies slightly more than halfway between Earth's center and surface.

The *P* and *S* waves caused by an earthquake do not travel in straight lines or at constant speed within Earth. The waves may reflect off boundaries between layers of different density, or they may be refracted as they pass through a boundary, and their speed is greater in regions of high temperature and density. Geoscientists can use the arrival times of reflected and refracted seismic waves from distant earthquakes to construct a detailed model of Earth's interior. Such studies confirm that the interior consists of three parts: a central core, a thick mantle, and a thin crust. Mathematical models based on the seismic data indicate that the core is hot (about 6000 K), dense (about 14 g/cm^3), and composed of iron and nickel. The core of Earth is as hot as the gasses at the sun's surface, but the high pressure keeps the metal solid near the center of the core and liquid in its outer part.

Earth's magnetism gives you further evidence about the core. The presence of a magnetic field is a clue that part of Earth's core must be a liquid metal. Convection currents stir the liquid, and it also rotates as Earth rotates. Because of these motions, and because it is a very good conductor of electricity, the liquid outer core generates a magnetic field through the dynamo effect—a different version of the process that creates the sun's magnetic field (look back to Chapter 7). From traces of magnetic field retained by rocks that formed long ago, geologists conclude that Earth's magnetic field has reversed itself in an irregular pattern, on average a few times per million years. Occasional reversals, though poorly understood, seem to be a characteristic of the dynamo effect.

The paths of seismic waves in the mantle, the layer of dense rock that lies between the molten core and the crust, show that it is not molten, but it is not really solid either. Mantle material behaves like a **plastic,** a material with the properties of a solid but capable of flowing under pressure. The asphalt used in paving roads is a common example of a plastic. It shatters if struck with a sledgehammer, but it bends under the steady weight of a heavy truck. Just below Earth's crust, where the pressure is less than at greater depths, the mantle is most plastic.

Earth's Active Crust

Earth's crust is composed of low-density rock that floats on the mantle. The image of a rock floating may seem odd, but recall that the rock of the mantle is very dense. Also, just below the crust, the mantle rock tends to be highly plastic, so great sections of low-density crust do indeed float on the semiliquid mantle like great lily pads floating on a pond. The crust is thickest under the continents, up to 60 km thick, and thinnest under the oceans, where it is only about 10 km thick. Unlike the mantle, the crust is brittle and can break when it is stressed.

The motion of the crust and the erosive action of water make Earth's crust highly active. Read **The Active Earth** on pages 172–173 and notice three important points and six new terms:

1. *Plate tectonics,* the motion of crustal plates, produces much of the geological activity on Earth. Plates spreading apart can form *rift valleys,* or, on the ocean floor, *mid-ocean rises* where molten rock solidifies to form *basalt.* A plate sliding into a *subduction zone* can trigger volcanism, and the collision of plates can produce *folded mountain ranges.* Chains of volcanoes such as the Hawaiian Islands can result when a plate moves horizontally across a hot spot.
2. Notice how the continents on Earth's surface have moved and changed over periods of hundreds of millions of years. A hundred million years is only 0.1 billion years, 1/46 of the age of Earth, so sections of Earth's crust are in rapid motion from the perspective of geologic time scales.
3. Most of the geological features you know—mountain ranges, the Grand Canyon, and even the familiar outline of the continents—are recent products of Earth's active surface.

Earth's surface is constantly renewed. The oldest rocks on Earth, small crystals of the mineral zircon from western Australia, are 4.4 billion years old. Most of the crust is much younger than that. Most of the mountains and valleys you see around you are no more than a few tens of millions of years old.

You can see that Earth's dramatic geology is dominated by two processes. Heat rising from the interior drives plate tectonics. Just below the thin crust of solid rock lies a semimolten layer with motions that can rip the crust to fragments, push the pieces about like rafts of wood on a pond, and sometimes crash them together to raise mountain ranges and grow continents. The second process modifying the crust is water. It falls as rain and snow and tears down mountains, erodes river valleys, and washes soil and rock into the sea. Tectonics builds up mountains and continents, and then erosion rips them down.

Earth's Atmosphere

You can't tell the story of Earth without mentioning its atmosphere. Not only is it necessary for life, but it is also intimately related to the crust. It affects the surface through erosion by wind and water, and in turn the chemistry of Earth's surface affects the composition of the atmosphere.

A modern understanding of planet building is that Earth formed so rapidly that it was substantially heated by the impacts of infalling material, as well as by radioactive decay, so the young Earth would have been highly volcanically active. When a

volcano erupts, 50 to 80 percent of the gas released is water vapor. The rest is mostly carbon dioxide, nitrogen, and smaller amounts of sulfur gases such as hydrogen sulfide—the rotten-egg gas that you smell if you visit geothermal pools and geysers such as those at Yellowstone National Park. If Earth's surface was molten as it formed, then outgassing would have been continuous, and Earth's original **primary atmosphere** would have been rich in water vapor, carbon dioxide, and nitrogen.

Astronomers also have suspected that some of the abundant water on Earth arrived late in the planet formation process as a bombardment of volatile-rich comets. Spectroscopic studies of comets revealed that some comets such as Comet LINEAR, which broke up in 1999 as it passed near the sun, have ratios of deuterium to hydrogen matching the ratio in the water on Earth, while other comets do not. Details of the origin of Earth's atmosphere and oceans are yet to be fully understood and are subjects of current research.

The atmosphere you breathe now is a **secondary atmosphere** produced later in Earth's history by chemical reactions between the atmosphere, crustal rocks, and surface water, and by green plants producing oxygen. Carbon dioxide is easily soluble in water—which is why carbonated beverages are so easy to manufacture—so the early oceans began to absorb atmospheric carbon dioxide. Once in solution, the carbon dioxide reacted with dissolved substances in the seawater to form silicon dioxide, limestone, and other mineral sediments in the ocean floor, freeing the seawater to absorb more carbon dioxide. Thanks to those chemical reactions in the oceans, most of the carbon dioxide was transferred from Earth's atmosphere to seafloor sediments.

When Earth was young, its atmosphere had no free oxygen, that is, oxygen not combined with other elements. Oxygen is very reactive and quickly forms oxides in the soil or combines with iron and other substances dissolved in water. Only the action of plant life keeps a steady supply of oxygen in Earth's atmosphere via photosynthesis, which makes energy for plants by absorbing carbon dioxide and releasing oxygen. Beginning about 2 to 2.5 billion years ago, photosynthetic plants in the oceans had multiplied to the point where they made oxygen at a rate faster than chemical reactions could remove it from the atmosphere. After that time, atmospheric oxygen increased rapidly (this topic will be discussed again in Chapter 20). It is a **Common Misconception** that there is life on Earth because of oxygen. The truth is exactly the opposite: There is oxygen in Earth's atmosphere because of life. Most lifeforms on Earth do not need oxygen (except the minority of creatures that are animals, including us), and some are even poisoned by it.

Once the atmosphere began to contain abundant ordinary oxygen (O_2), Earth's lower atmosphere became protected from solar ultraviolet radiation by an **ozone layer** about 15 to 30 km above the surface. Ozone molecules are very good at absorbing ultraviolet photons. An ozone molecule consists of three oxygen atoms linked together (O_3). Before there was oxygen, an ozone layer could not form, and the sun's ultraviolet radiation was able to penetrate deep into the atmosphere and break up weaker molecules such as water (H_2O). The hydrogen from the water then escaped to space, and the oxygen formed oxides in the crust. Earth's atmosphere could not reach its present stable composition until it was protected by an ozone layer, and that required oxygen.

Human Effects on the Atmosphere and Climate

You can live on Earth's surface because of Earth's atmosphere, but modern civilization is altering Earth's atmosphere in at least two serious ways, by adding carbon dioxide (CO_2) and by destroying ozone.

The concentration of CO_2 in Earth's atmosphere is important because CO_2 can trap heat in a process called the **greenhouse effect** (■ Figure 9-4). When sunlight shines through the glass roof of a greenhouse, it heats the benches and plants inside. The warmed interior radiates infrared radiation, but the glass is opaque to infrared. Warm air in the greenhouse cannot mix with cooler air outside, so heat is trapped within the greenhouse, and the temperature climbs until the glass itself grows warm enough to radiate heat away as fast as the sunlight enters. This is the same process that heats a car when it is parked in the sun with the windows rolled up.

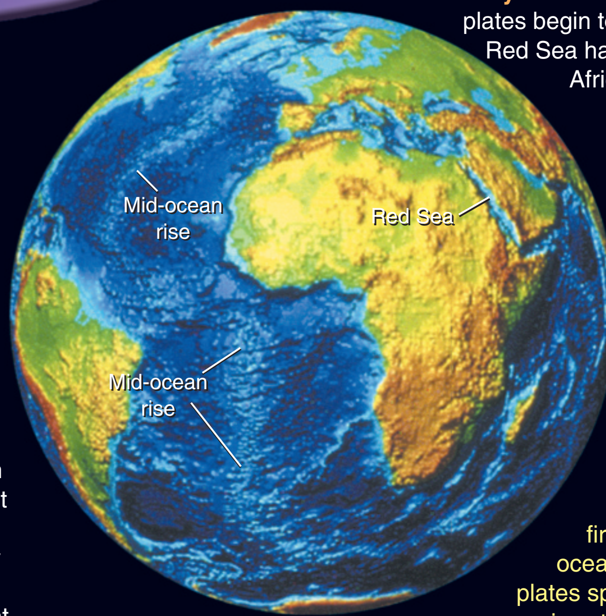
Earth's atmosphere is transparent to sunlight, and when the ground absorbs the sunlight, it grows warmer and radiates at infrared wavelengths. However, CO_2 makes the atmosphere less transparent to infrared radiation, so infrared radiation from the warm surface is absorbed by the atmosphere and cannot escape back into space. That traps heat and makes Earth warmer (Figure 9-4b). CO_2 is not the only greenhouse gas. Water vapor, methane, and other gases also help warm Earth, but CO_2 is the most important. It is a **Common Misconception** that the greenhouse effect is only bad. Evidence indicates that Earth has had a greenhouse effect for its entire history. Without the greenhouse effect, Earth presently would be at least 30°C (54°F) colder and uninhabitable for most living things, including humans.

For 4 billion years, natural processes on Earth have removed CO_2 from the atmosphere and buried the carbon in the form of limestone, coal, oil, and natural gas. Since the beginning of the Industrial Revolution in the late 18th century, humans have been digging up lots of carbon-rich fuels, burning them to get energy and releasing CO_2 back into the atmosphere more rapidly than it can be naturally removed, thereby increasing the intensity of the greenhouse effect. The increased concentration of CO_2 is increasing the greenhouse effect and warming Earth in what is known as **global warming.**

It is a **Common Misconception** that human output of CO_2 is minor compared to natural sources such as volcanoes. Careful measurements of carbon isotope ratios and relative amounts of CO_2 versus O_2 in the atmosphere show that the CO_2 added to the atmosphere since the year 1800 is mostly or entirely due to human burning of fossil fuels. Estimates are that the amount of CO_2 in Earth's atmosphere could double during the 21st century.

1 Our world is an astonishingly active planet. Not only is it rich in water and therefore subject to rapid erosion, but its crust is divided into moving sections called plates. Where plates spread apart, lava wells up to form new crust; where plates push against each other, they crumple the crust to form mountains. Where one plate slides over another, you see volcanism. This process is called **plate tectonics,** referring to the Greek word for "builder." (An architect is literally an arch builder.)

A **rift valley** forms where continental plates begin to pull apart. The Red Sea has formed where Africa has begun to pull away from the Arabian peninsula.



National Geophysical Data Center

1a Evidence of plate tectonics was first found in ocean floors, where plates spread apart and magma rises to form **mid-ocean rises** made of rock called **basalt,** a rock typical of solidified lava. Radioactive dating shows that the basalt is younger near the mid-ocean rise. Also, the ocean floor carries less sediment near the mid-ocean rise. As Earth's magnetic field reverses back and forth, it is recorded in the magnetic fields frozen into the basalt. This produces a magnetic pattern in the basalt that shows that the seafloor is spreading away from the mid-ocean rise.

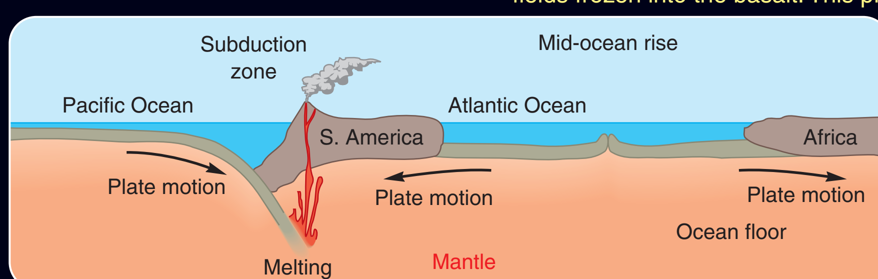
A typical view of planet Earth

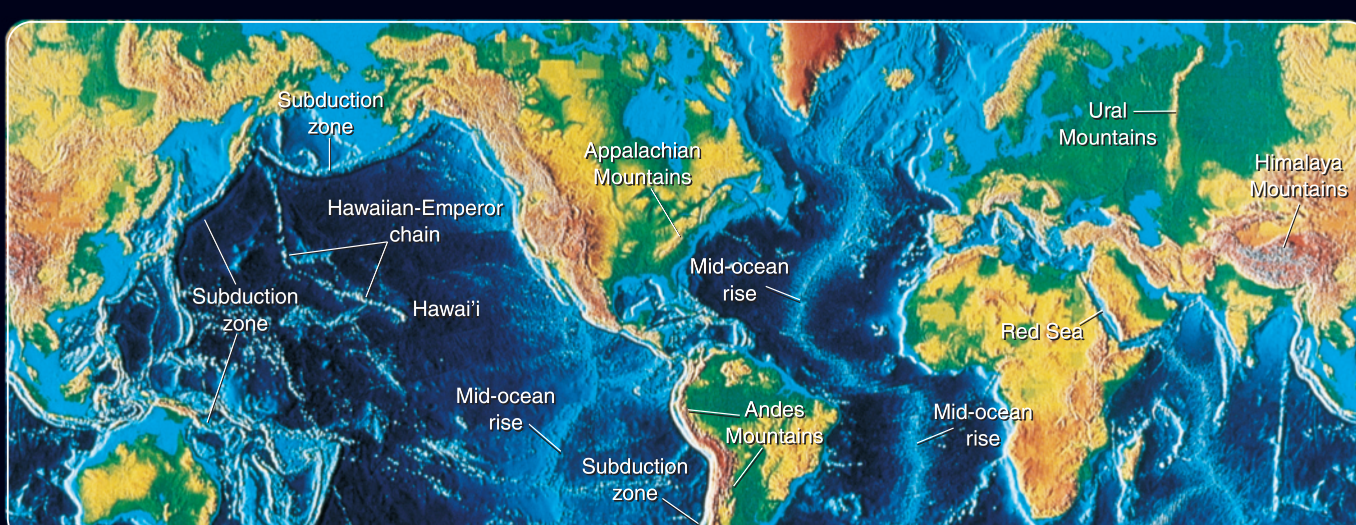
William K. Hartmann/Planetary Science Institute

Courtesy Janet Seeds

Mountains are common on Earth, but they erode away rapidly because of the abundant water.

1b A **subduction zone** is a deep trench where one plate slides under another. Melting releases low-density magma that rises to form volcanoes such as those along the northwest coast of North America, including Mt. St. Helens.

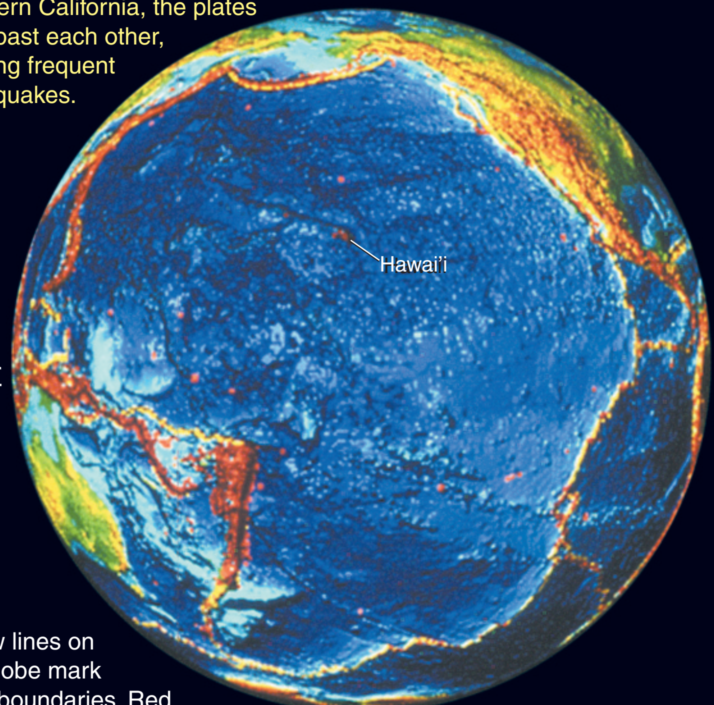




National Geophysical Data Center

1c Hot spots caused by rising magma in the mantle can poke through a plate and cause volcanism such as that in Hawai'i. As the Pacific plate has moved northwestward, the hot spot has punched through to form a chain of volcanic islands, now mostly worn below sea level. **Folded mountain ranges** can form where plates push against each other. For example, the Ural Mountains lie between Europe and Asia, and the Himalaya Mountains are formed by India pushing north into Asia. The Appalachian Mountains are the remains of a mountain range thrust up when North America was pushed against Europe and Africa.

1d The floor of the Pacific Ocean is sliding into subduction zones in many places around its perimeter. This pushes up mountains such as the Andes and triggers earthquakes and active volcanism all around the Pacific in what is called the Ring of Fire. In places such as southern California, the plates slide past each other, causing frequent earthquakes.



National Geophysical Data Center

Yellow lines on this globe mark plate boundaries. Red dots mark earthquakes since 1980. Earthquakes within the plate, such as those at Hawai'i, are related to volcanism over hot spots in the mantle.

Continental Drift

Not long ago, Earth's continents came together to form one continent dubbed Pangea by geoscientists.

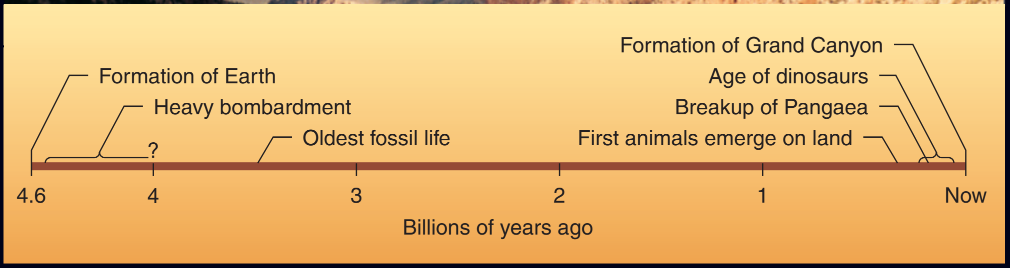
Pangaea broke into a northern and a southern continent.

Notice India moving north toward Asia.

The continents are still drifting on the highly plastic upper mantle.

2 The floor of the Atlantic Ocean is not being subducted. It is locked to the continents and is pushing North and South America away from Europe and Africa at about 3 cm per year, a motion called *continental drift*. Scientists can measure the motion by, for example, comparing the time for laser flashes sent from European and American observatories to bounce off mirrors left on the moon by the Apollo astronauts. Roughly 200 million years ago, North and South America were joined to Europe and Africa. Evidence of that lies in similar fossils and similar rocks and minerals found in the matching parts of the continents. Notice how North and South America fit against Europe and Africa like a puzzle.

M. Seeds



3 Plate tectonics pushes up mountain ranges and causes bulges in the crust, and water erosion wears the rock away. The Colorado River began cutting the Grand Canyon only about 10 million years ago when the Colorado plateau warped upward under the pressure of moving plates. That sounds like a long time ago, but it is only 0.01 billion years. A mile down, at the bottom of the canyon, lie rocks 0.57 billion years old, the roots of an earlier mountain range that stood as high as the Himalayas. It was pushed up, worn away to nothing, and covered with sediment long ago. Many of the geological features we know on Earth have been produced by relatively recent events.

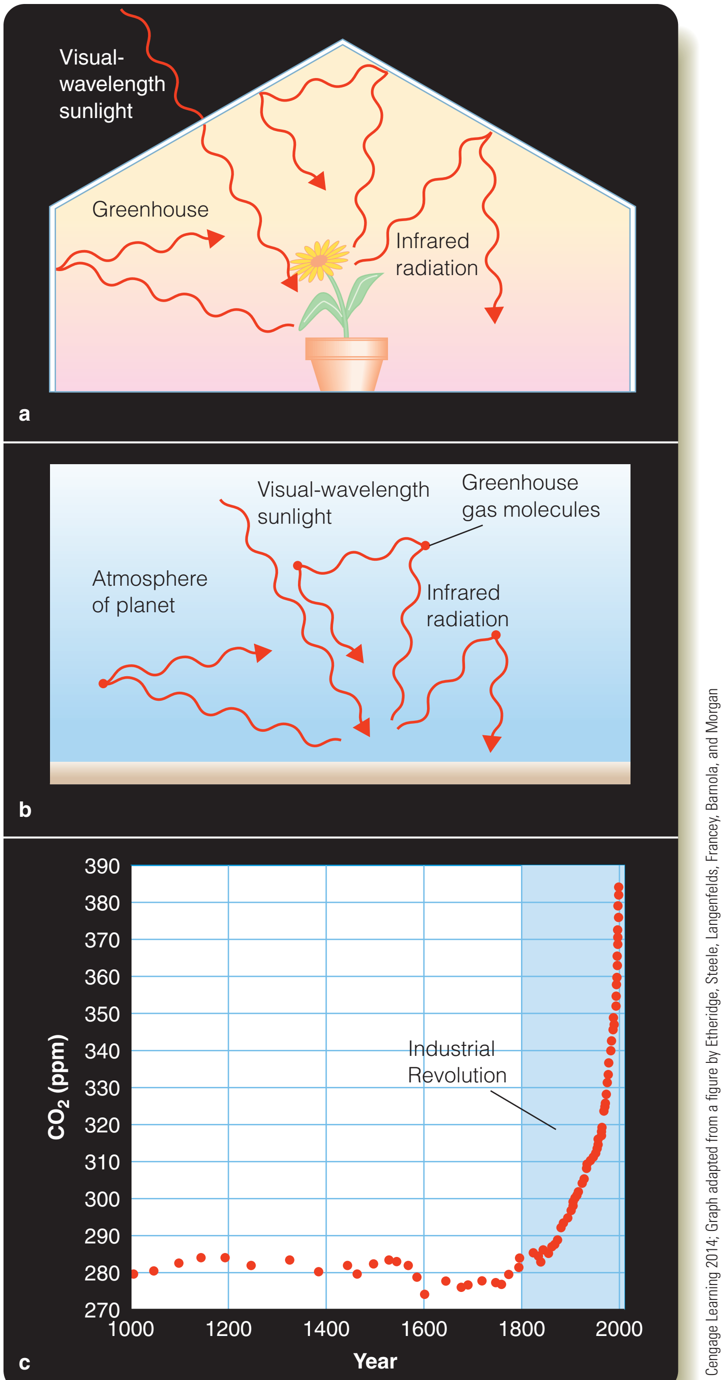


Figure 9-4

The greenhouse effect: (a) Visual-wavelength sunlight can enter a greenhouse and heat its contents, but the longer-wavelength infrared radiation cannot get out. (b) The same process can heat a planet's surface if its atmosphere contains greenhouse gases such as CO_2. (c) The concentration of CO_2 in Earth's atmosphere as measured in Antarctic ice cores remained roughly constant for thousands of years until the beginning of the Industrial Revolution around the year 1800. Since then it has increased by more than 30 percent. Evidence from proportions of carbon isotopes and oxygen in the atmosphere proves that most of the added CO_2 is the result of burning fossil fuels.

Studies of the growth rings in very old trees show that the average Earth climate had been cooling for most of the last 1000 years, but the 20th century reversed that trend with a rise of 0.56 to 0.92°C (1.01 to 1.66°F). Mountain glaciers have melted back dramatically since the 19th century. Measurements show that polar ice in the form of permafrost, ice shelves, and ice on the open Arctic Ocean is melting.

Yet another **Common Misconception** is that the observed warming of the Earth is due to natural causes rather than the greenhouse effect. As scientists began to study early indications that Earth's global climate is warming and that human-produced CO_2 might be a major cause, they reacted with professional skepticism (**How Do We Know? 9-2**). They asked questions such as, "Is global warming really happening?" and, "If the warming is real, is some mechanism other than human activity the main cause?" Experiments were conducted, calculations performed, and models created to answer those types of questions. Such research, carried out over decades, revealed that regular and predictable changes in Earth's axis inclination and orientation and in the shape of its orbit, called Milankovitch cycles (Chapter 3), currently would be driving Earth's climate toward lower, not higher, temperatures. Also, observations by space probes indicate that the sun's luminosity, averaged over its activity cycles, has been constant for decades or has decreased slightly. The observed warming must be strong to be occurring in the face of opposing astronomical effects.

The amount of warming to expect in the future is difficult to predict because Earth's climate is critically sensitive to a number of different factors, not just the abundance of greenhouse gases. For example, a slight warming should increase water vapor in the atmosphere, and although water vapor is another greenhouse gas that would enhance the warming, increased water vapor might cause added cloud cover, increasing Earth's **albedo,** the fraction of light it receives that is reflected back to space. An increased albedo would tend to reduce the warming. Also, even small changes in temperature can alter circulation patterns in the atmosphere and in the oceans, and the consequences of such changes are very difficult to model. The situation is complex, but models that best track the warming that has already happened point to substantial continued warming.

There is no doubt that civilization is warming Earth through an enhanced greenhouse effect, but a remedy is difficult to imagine. Reducing the amount of CO_2 and other greenhouse gases released to the atmosphere is difficult because modern society depends on burning fossil fuels for energy. Political, business, and economic leaders may argue that the issue is uncertain, but all around the world scientists of stature have reached agreement: Global warming is real, is driven by human activity, and will change Earth. What humanity can or will do about it is uncertain.

Human influences on Earth's atmosphere go beyond the greenhouse effect. Our modern industrial civilization is also reducing ozone in Earth's atmosphere. Many people have a **Common Misconception** that ozone is bad because they hear it

How Do We Know? 9-2

Scientists: Courteous Skeptics

What does it mean to be skeptical, yet also open to new ideas? "Scientists are just a bunch of skeptics who don't believe in anything." That is a **Common Misconception** among people who don't understand the methods and goals of science. Yes, scientists are skeptical about new ideas and discoveries, but they do hold strong beliefs about how nature works. Scientists are skeptical not because they want to disprove everything but because they are searching for the truth and want to be sure that a new description of nature is reliable before it is accepted.

Another **Common Misconception** is that scientists automatically accept the work of other scientists. On the contrary, scientists skeptically question every aspect of a new discovery. They may wonder if another scientist's instruments were properly calibrated or whether the scientist's mathematical models are correct. Other scientists will want to repeat the work themselves using their own instruments to see if they can obtain the same results. Every observation is tested, every discovery is confirmed, and only an idea that survives many of these tests begins to be accepted as a scientific truth.

Scientists are prepared for this kind of treatment at the hands of other scientists. In fact, they expect it. Among scientists it is not bad manners to say, "Really, how do you know that?" or "Why do you think that?" or "Show me the evidence!" It is not just new or surprising claims that are subject to such scrutiny. Although there was some evidence that global warming is occurring and that human-produced CO_2 is a major cause, scientists were professionally skeptical about those hypotheses. This was not because they thought the early observations and models were obviously flawed, or because they had strong personal skepticism, but because that is how science works.

The goal of science is to tell stories about nature. Some people use the phrase "telling a story" to describe someone who is telling a fib. But the stories that scientists tell are the opposite; perhaps you could call them "antifibs" because they are as true as scientists can make them. Skepticism eliminates stories with logical errors, flawed observations, or misunderstood evidence and eventually leaves only the stories that best describe nature.

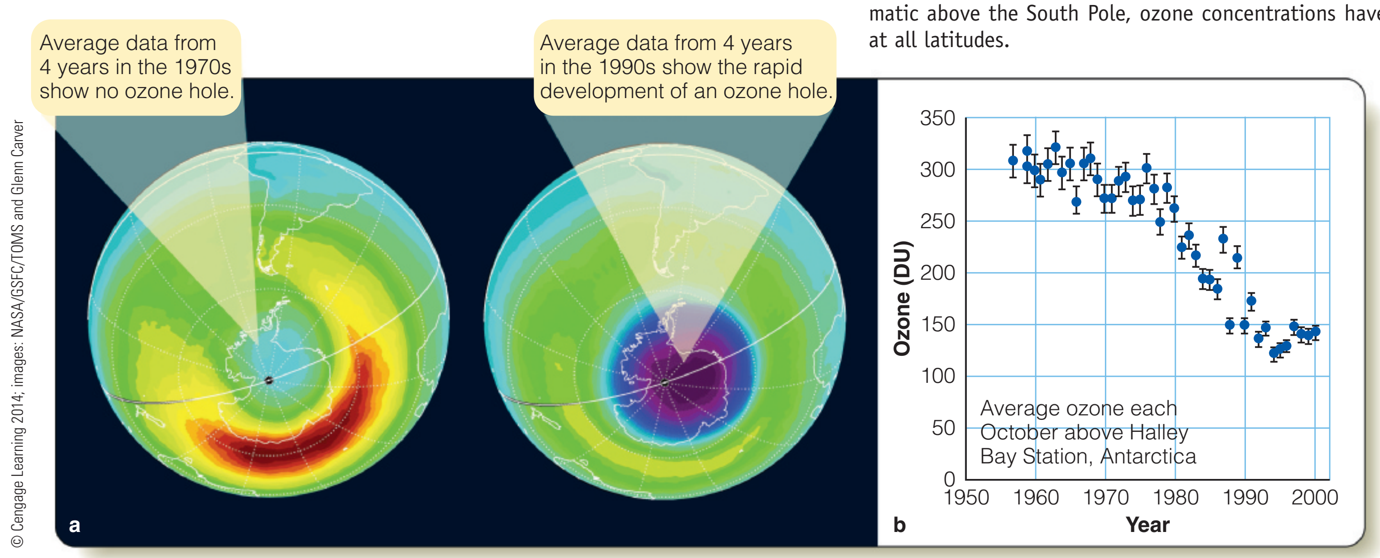
Skepticism is not a refusal to hold beliefs. Rather, it is a way for scientists to find and keep those natural principles that are worthy of belief.

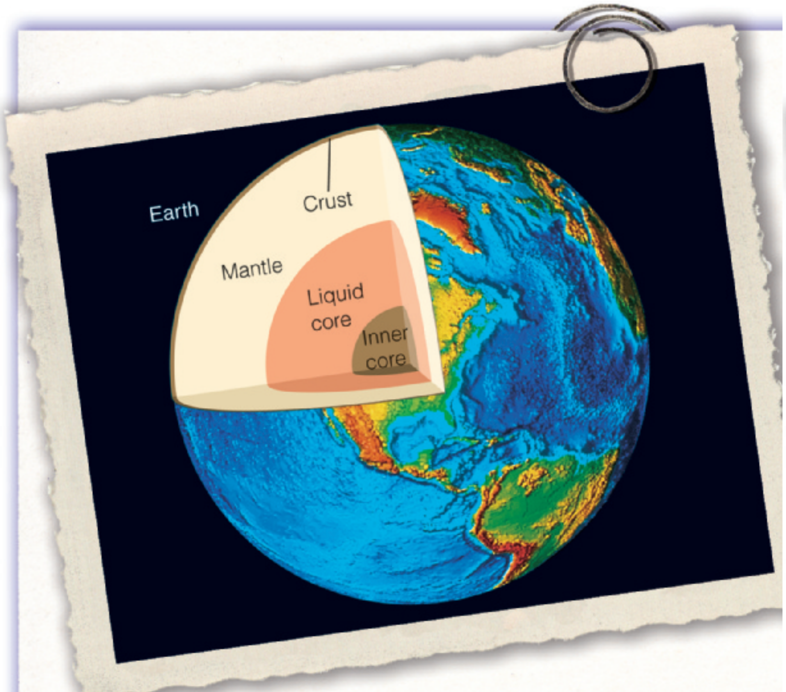
mentioned as a pollutant of city air, produced by auto emissions. Breathing ozone is bad for you, but, as you learned earlier in this chapter, the ozone layer in the upper atmosphere protects the lower atmosphere and Earth's surface from harmful solar UV photons. Ozone (O_3) is an unstable molecule. Certain chemicals called chlorofluorocarbons (CFCs), used for refrigeration, air-conditioning, and some industrial processes, can easily destroy ozone. As these CFCs escape into the atmosphere, they become mixed into the ozone layer and convert the ozone back into normal oxygen molecules. Ordinary oxygen does not block ultraviolet radiation, so depleting the ozone layer causes an increase in ultraviolet radiation at Earth's surface. In small doses, ultraviolet radiation can produce a suntan, but in larger doses it can cause skin cancers.

The ozone layer is especially sensitive to CFCs over the Antarctic and Arctic regions because of their temperatures (■ Figure 9-5). Starting in the late 1970s, the ozone concentration fell significantly over the Antarctic, and a hole in the ozone layer

■ Figure 9-5

(a) Satellite observations of ozone concentrations over Antarctica are shown here as red for highest concentration and violet for lowest. Since the 1970s, a hole in the ozone layer has developed over the South Pole. (b) Although ozone depletion is most dramatic above the South Pole, ozone concentrations have declined at all latitudes.





© Cengage Learning 2014; image: NGDC

Earth's surface is marked by high continents and low seafloors. The crust is only 10 to 60 km thick. Interior to that are a thick mantle, liquid outer core, and solid inner core.

Celestial Profile 2: The Earth

Motion:

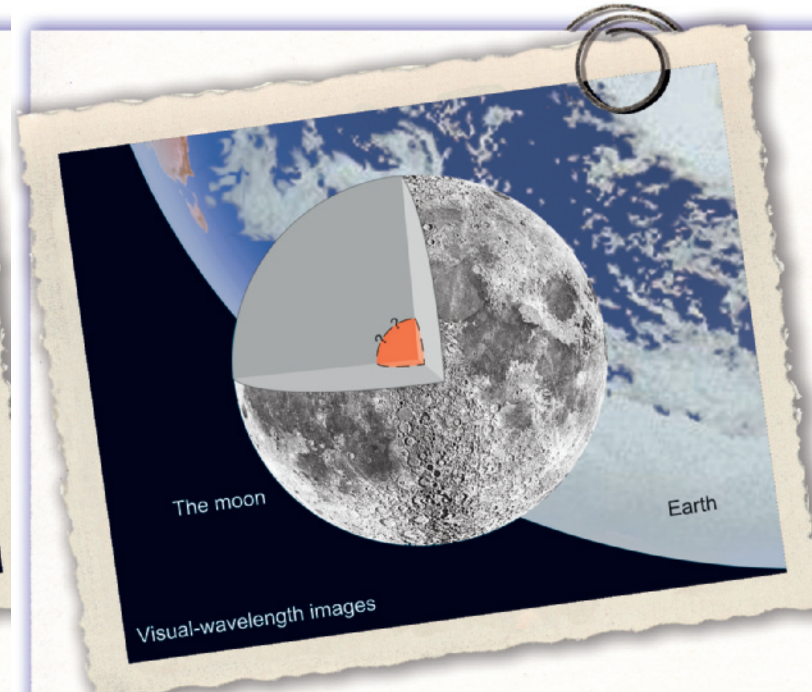
Average distance from the sun	1.00 AU (1.50×10^8 km)
Eccentricity of orbit	0.017
Inclination of orbit to ecliptic	0° (by definition)
Average orbital velocity	29.8 km/s
Orbital period	1.0000 y (365.26 days)
Period of rotation	24.00 h (with respect to the sun)
Period of rotation	23.93 h (with respect to the stars)
Inclination of equator to orbit	23.4°

Characteristics:

Equatorial diameter	1.28×10^4 km
Mass	5.97×10^{24} kg
Average density	5.52 g/cm^3 (4.07 g/cm^3 uncompressed)
Surface gravity	1.00 Earth gravity
Escape velocity	11.2 km/s
Surface temperature	−90° to 60°C (−130° to 140°F)
Average albedo	0.31
Oblateness	0.0034

Personality Point:

The modern English word *Earth* comes from Old English *eorthe* and ultimately from the Indo-European root *er-*, meaning ground or dirt. *Terra* comes from the Roman goddess of fertility and growth; thus, *Terra Mater*, Mother Earth.



© Cengage Learning 2014; image: NASA

Earth's moon has about a quarter the diameter of Earth. Its low density indicates that it does not contain much iron. The size of its core, if any, and the amount of remaining heat are unknown.

Celestial Profile 3: The Moon

Motion:

Average distance from Earth	3.84×10^5 km (center to center)
Eccentricity of orbit	0.055
Inclination of orbit to ecliptic	5.1°
Average orbital velocity	1.02 km/s
Orbital period (sidereal)	27.3 d
Orbital period (synodic)	29.5 d
Inclination of equator to orbit	6.7°

Characteristics:

Equatorial diameter	3.48×10^3 km ($0.273\ D_\oplus$)
Mass	7.35×10^{22} kg ($0.0123\ M_\oplus$)
Average density	3.35 g/cm^3 (3.3 g/cm^3 uncompressed)
Surface gravity	0.17 Earth gravity
Escape velocity	2.4 km/s ($0.21\ V_\oplus$)
Surface temperature	−170° to 130°C (−275° to 265°F)
Average albedo	0.12
Oblateness	0

Personality Point:

Lunar superstitions are common. The words *lunatic* and *lunacy* come from *luna*, the moon. Someone who is *moonstruck* is supposed to be a bit nutty. Because the moon affects the ocean tides, many superstitions link the moon to water, to weather, and to women's cycle of fertility. According to legend, moonlight is supposed to be harmful to unborn children, but on the plus side, moonlight rituals are said to remove warts.

developed over the continent each October at the time of the Antarctic spring. Satellite and ground-based measurements showed the same thing beginning to happen at far northern latitudes, with the amount of ultraviolet radiation reaching the ground increasing. Fortunately, as a result of these warnings, international agreements banned most uses of CFCs, and the trend of ozone hole expansion seems to have slowed and may be reversing.

There is yet another **Common Misconception** that global warming and ozone depletion are two names for the same thing. Take careful note that the ozone hole is a second Earth environmental issue that is basically separate from global warming. The CO_2 and ozone problems in Earth's atmosphere are paralleled on Venus and Mars. When you study Venus in Chapter 10, you will discover a runaway greenhouse effect that has made the surface of the planet hot enough to melt lead. Also in Chapter 10, you will learn that Mars has an atmosphere without an ozone layer. A few minutes of sunbathing on Mars would kill you. Once again, you can learn more about your own planet by studying exaggerated conditions on other planets.

9-3 The Moon

If you had been one of the first two people from Earth to step onto the airless surface of the moon, what would you have said? Neil Armstrong responded to the historic significance of the moment by saying, "That's one small step for [a] man, one giant leap for mankind." Buzz Aldrin was second, and he responded to the moon itself, saying, "Beautiful, beautiful. Magnificent desolation!" In other words, although it is desolate, the moon has its own kind of beauty. Many planets in the universe probably look like Earth's moon, and astronauts may someday walk on such worlds and compare them with Earth's moon.

Only twelve people have stood on the moon, but planetary astronomers know it well. The photographs, measurements, and samples brought back to Earth paint a picture of an ancient battered crust, and a world created by a planetary catastrophe.

Lunar Geology

The surface of the moon is divided into two dramatically different kinds of terrain. The dark gray areas visible from Earth by naked eye are the smooth lunar lowlands, which, using the Latin word for *seas,* earlier astronomers named **maria** (plural of **mare,** which is pronounced *mah-ray*). You can also see the comparatively bright, rugged, and heavily cratered lunar highlands.

The color of moon rocks is dark gray, but Earth's moon looks quite bright in the night sky. In fact, the average albedo of Earth's moon, the fraction of the light that it reflects, is only 0.12 (**Celestial Profile 3**, p. 176). In other words, the moon reflects only 12 percent of the light of the sunlight that hits it. In comparison, Earth, thanks mostly to its bright clouds, has an average albedo of 0.31. The moon looks bright only in contrast to the night sky. In reality it is a dark gray world.

Wherever you look on the moon, you find craters. These craters look quite dramatic near the **terminator,** the name for the boundary between daylight and darkness on the moon where shadows are long. As you have already learned, the highlands are heavily marked by craters whereas the smooth lowlands contain relatively few craters.

Planetary scientists now know that the craters on the moon were formed by the impact of meteorites. Study **Impact Cratering** on pages 178–179 and notice three important points:

1. Impact craters have certain distinguishing characteristics, such as their shape and the *ejecta, rays,* and *secondary craters* around them.
2. Lunar impact craters range from tiny pits formed by *micrometeorites* to giant *multiringed basins.*
3. Most of the craters on the moon are old; they were formed long ago when the solar system was young.

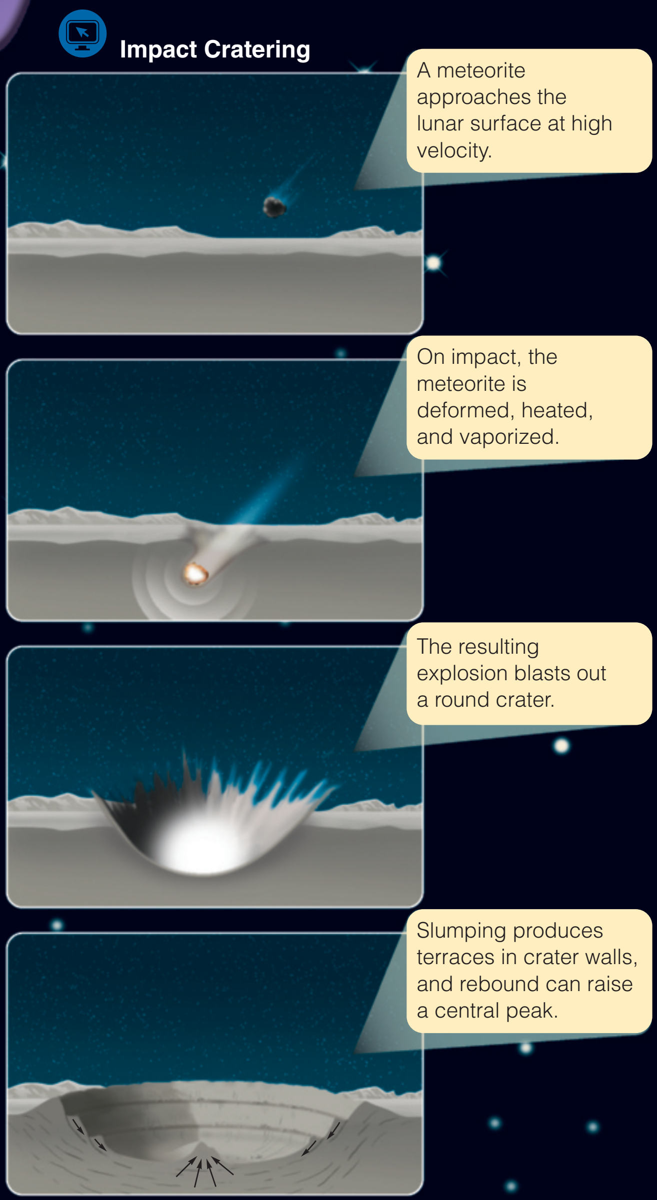
Twelve Apollo astronauts visited the lunar lowlands and highlands between 1969 and 1972 (■ Figure 9-6). Most of the rocks they found were typical of hardened lava, and some were **vesicular basalt,** which contains holes formed by bubbles in the molten rock (■ Figure 9-7). These bubbles are made when lava flows out onto the surface, and the lower pressure allows gases dissolved in the molten rock to expand and form bubbles. The same thing happens when you open a bottle of carbonated beverage and bubbles form. The presence of vesicular basalts shows that much of the surface of Earth's moon has been covered by successive lava flows, and the dark flat plains of the lunar lowlands, the maria, are actually solidified ancient lava. The highlands, in contrast, are composed of rock containing minerals that have low density and would be among the first to solidify and float to the top of molten rock. For example, the highlands are rich in **anorthosite,** a light-colored and low-density rock that contributes to the highlands' bright contrast with the dark lowlands.

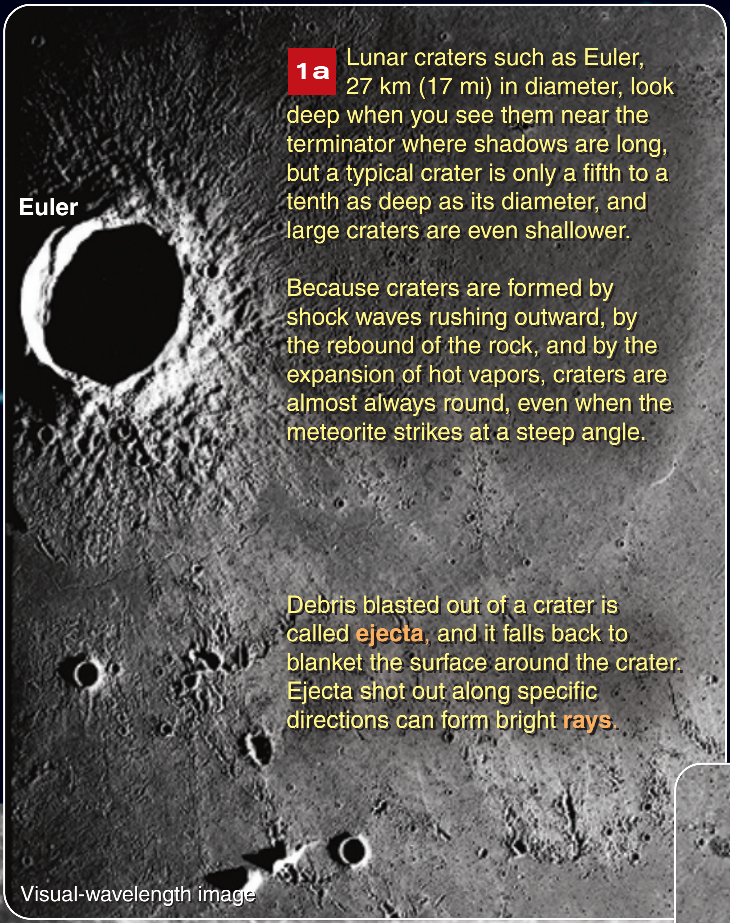
Many of the rocks all over the moon are **breccias,** rocks made up of fragments of broken rock cemented together under pressure (Figure 9-7). The breccias show how extensively the lunar surface has been pounded by meteorites. Nowhere did the astronauts find what could be called bedrock; the entire surface of Earth's moon is fractured by meteorite impacts. Moreover, as the astronauts bobbed across the lunar surface under its low gravity, their boots kicked up the powdery dust. This lunar dust is produced by the continuous bombardment of the lunar surface by tiny meteorites that slowly grind exposed rocks into fine gray grit with a consistency like talcum powder. The *LCROSS (Lunar Crater Observation and Sensing Satellite)* mission discovered that, at least in one permanently shadowed region near the moon's south pole, there is a significant amount of water ice under the surface dust layer.

Impact Cratering

1 The craters that cover the moon and many other bodies in the solar system were produced by the high-speed impact of meteorites of all sizes. Meteorites striking the moon travel 10 to 70 km/s and can hit with the energy of many nuclear bombs.

A meteorite striking the moon's surface can deliver tremendous energy and can produce an impact crater 10 or more times larger in diameter than the meteorite. The vertical scale is exaggerated at right for clarity.





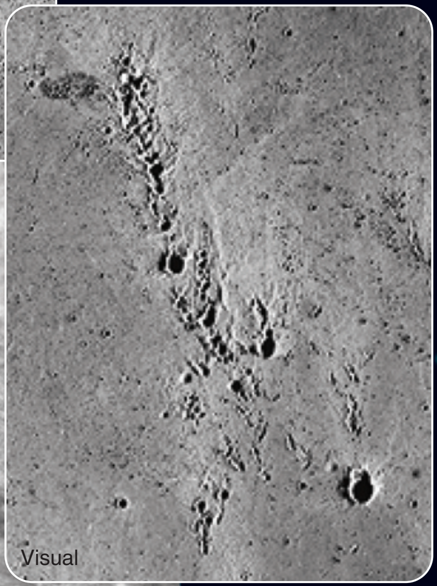
1a Lunar craters such as Euler, 27 km (17 mi) in diameter, look deep when you see them near the terminator where shadows are long, but a typical crater is only a fifth to a tenth as deep as its diameter, and large craters are even shallower.

Because craters are formed by shock waves rushing outward, by the rebound of the rock, and by the expansion of hot vapors, craters are almost always round, even when the meteorite strikes at a steep angle.

Debris blasted out of a crater is called **ejecta**, and it falls back to blanket the surface around the crater. Ejecta shot out along specific directions can form bright **rays**.

Visual-wavelength image

NASA

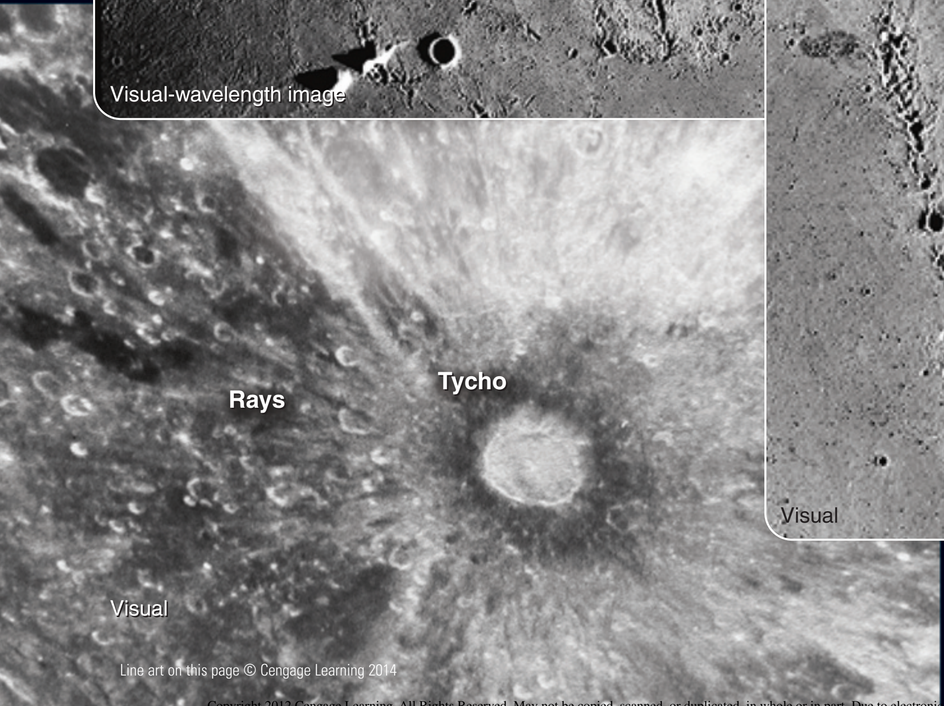


Visual

NASA

1b Rock ejected from distant impacts can fall back to the surface and form smaller craters called **secondary craters.** The chain of craters here is a 45-km-long chain of secondary craters produced by ejecta from the large crater Copernicus 200 km out of the frame to the lower right.

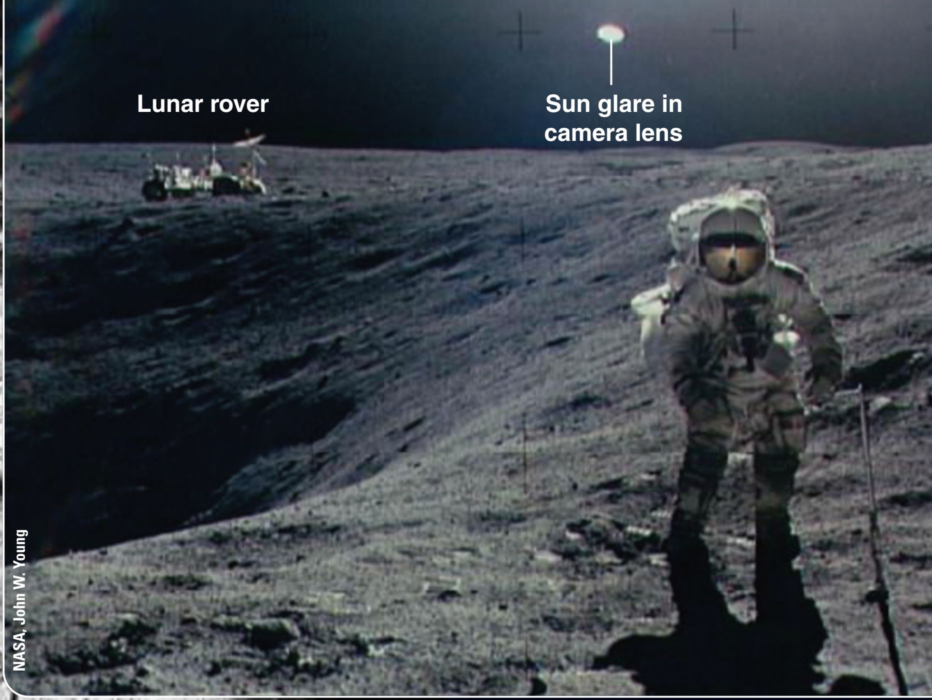
Bright ejecta blankets and rays gradually darken as sunlight alters minerals and small meteorites stir the dusty surface. Bright rays are signs of youth. Rays from the crater Tycho, perhaps only 100 million years old, extend halfway around the moon.



Visual

NASA

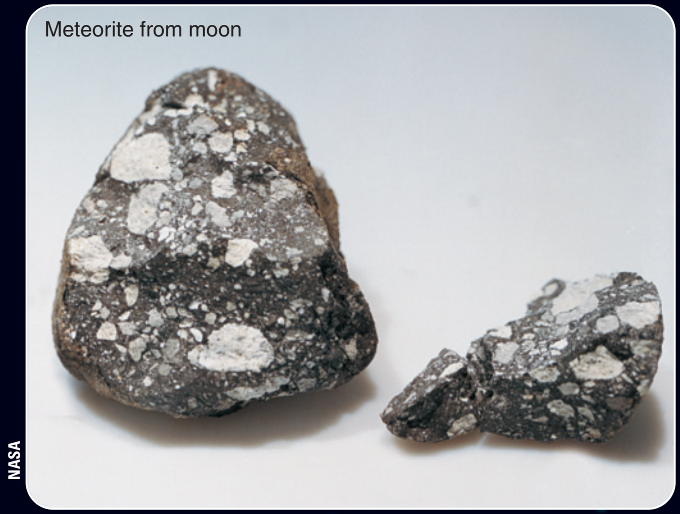
2 Plum Crater, 40 m (130 ft) in diameter, was visited by Apollo 16 astronauts. Note the many smaller craters visible. Lunar craters range from giant impact basins to tiny pits in rocks struck by **micrometeorites,** meteorites of microscopic size.



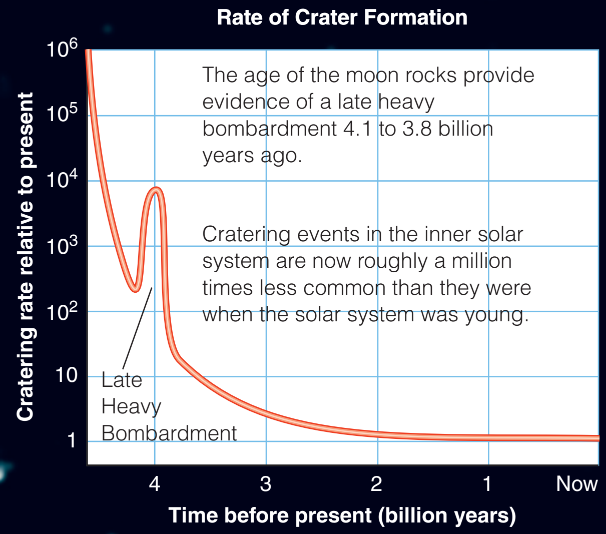
2a In larger craters, the deformation of the rock can form one or more inner rings concentric with the outer rim. The largest of these craters are called **multiringed basins.** In Mare Orientale on the west edge of the visible moon, the outermost ring is almost 900 km (550 mi) in diameter.

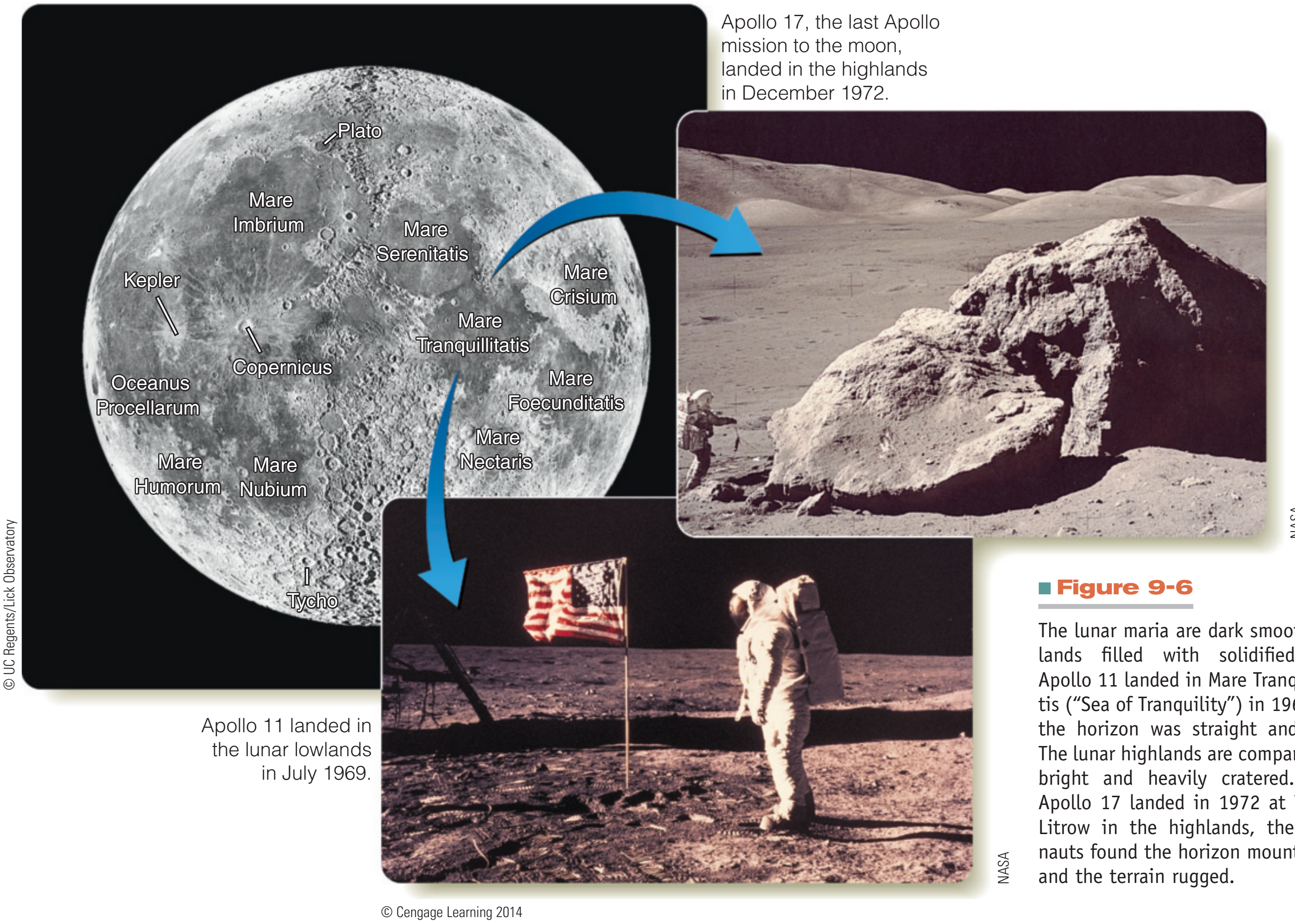
2b The energy of an impact can melt rock, some of which falls back into the crater and solidifies. When the moon was young, craters could also be flooded by lava welling up from below the crust.

A few meteorites found on Earth have been identified chemically as fragments of the moon's surface blasted into space by cratering impacts. The fragmented nature of these meteorites indicates that the moon's surface has been battered by impact craters.



3 Most of the craters on the moon were produced long ago when the solar system was filled with debris from planet building. As that debris was swept up, the cratering rate fell rapidly, shown schematically below.





Apollo 17, the last Apollo mission to the moon, landed in the highlands in December 1972.

Apollo 11 landed in the lunar lowlands in July 1969.

Figure 9-6

The lunar maria are dark smooth lowlands filled with solidified lava. Apollo 11 landed in Mare Tranquillitatis ("Sea of Tranquility") in 1969, and the horizon was straight and level. The lunar highlands are comparatively bright and heavily cratered. When Apollo 17 landed in 1972 at Taurus-Litrow in the highlands, the astronauts found the horizon mountainous and the terrain rugged.

Origin of Earth's Moon

Over the last two centuries, astronomers developed three different hypotheses for the origin of Earth's moon. The *fission hypothesis* proposed that the moon broke from a rapidly spinning young Earth. The *condensation hypothesis* suggested that Earth and its moon condensed from the same cloud of matter in the solar nebula. The *capture hypothesis* suggested that the moon formed elsewhere in the solar nebula and was later captured by Earth. Each of these ideas had problems and failed to survive comparison with all the evidence.

In the 1970s, a new hypothesis originated that combined some aspects of the three older hypotheses. The **large-impact hypothesis** proposes that the moon formed when a very large planetesimal, estimated to have been at least as massive as Mars, smashed into the proto-Earth. Such a collision would have ejected debris into space, and model calculations indicate that some of the debris would have settled into a disk in orbit around Earth. The material in the disk would then have quickly collected to form the moon (■ Figure 9-8).

This hypothesis explains several phenomena. If the collision occurred off-center, it would have spun the Earth–moon system rapidly and would thus explain the present high angular momentum. If the proto-Earth and impactor had each already differentiated, the ejected material that became the moon would have been mostly iron-poor mantle and crust, which would explain the moon's low density and iron-poor composition. Furthermore, the material would have lost its volatile components while it was in space, so the moon also would have formed lacking volatiles. Such an impact would have melted the proto-Earth, and the material falling together to form the moon would also have been heated hot enough to melt. This fits the evidence that the highland anorthosite in the moon's oldest rocks formed by differentiation of large quantities of molten material. The large-impact hypothesis survives comparison with the known evidence and is now considered likely to be correct.

History of Earth's Moon

The four-stage history of Earth's moon is dominated by a single fact that makes its unfolding noticeably different from Earth's history in the later stages. The moon is small, only one-fourth the diameter of Earth. Its escape velocity is low, so the moon has been unable to hold any atmosphere, cannot have surface water, and its interior cooled rapidly as its internal heat flowed outward into space. Small worlds have less heat and lose it more rapidly, so the moon's small size has been critical in determining its history.

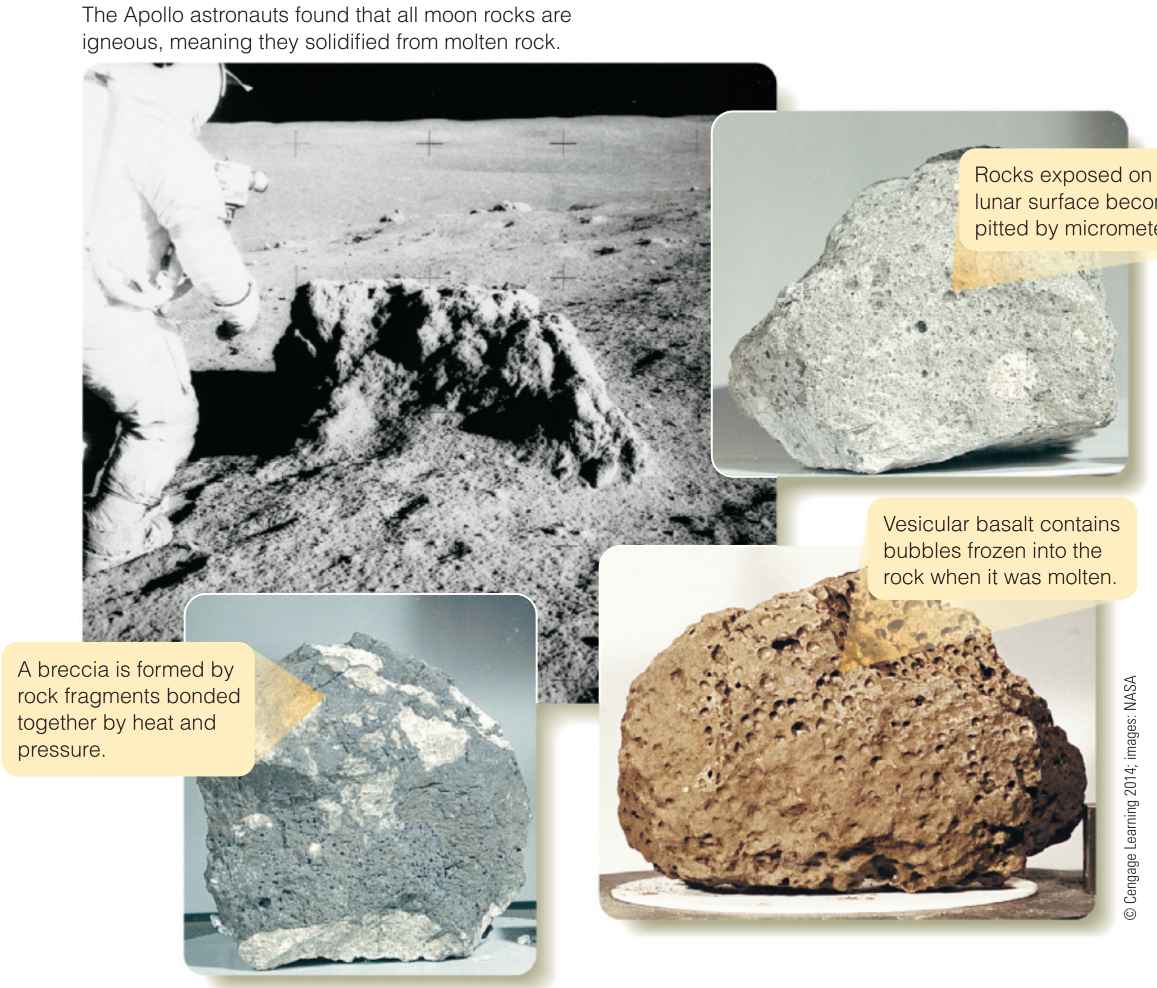


Figure 9-7

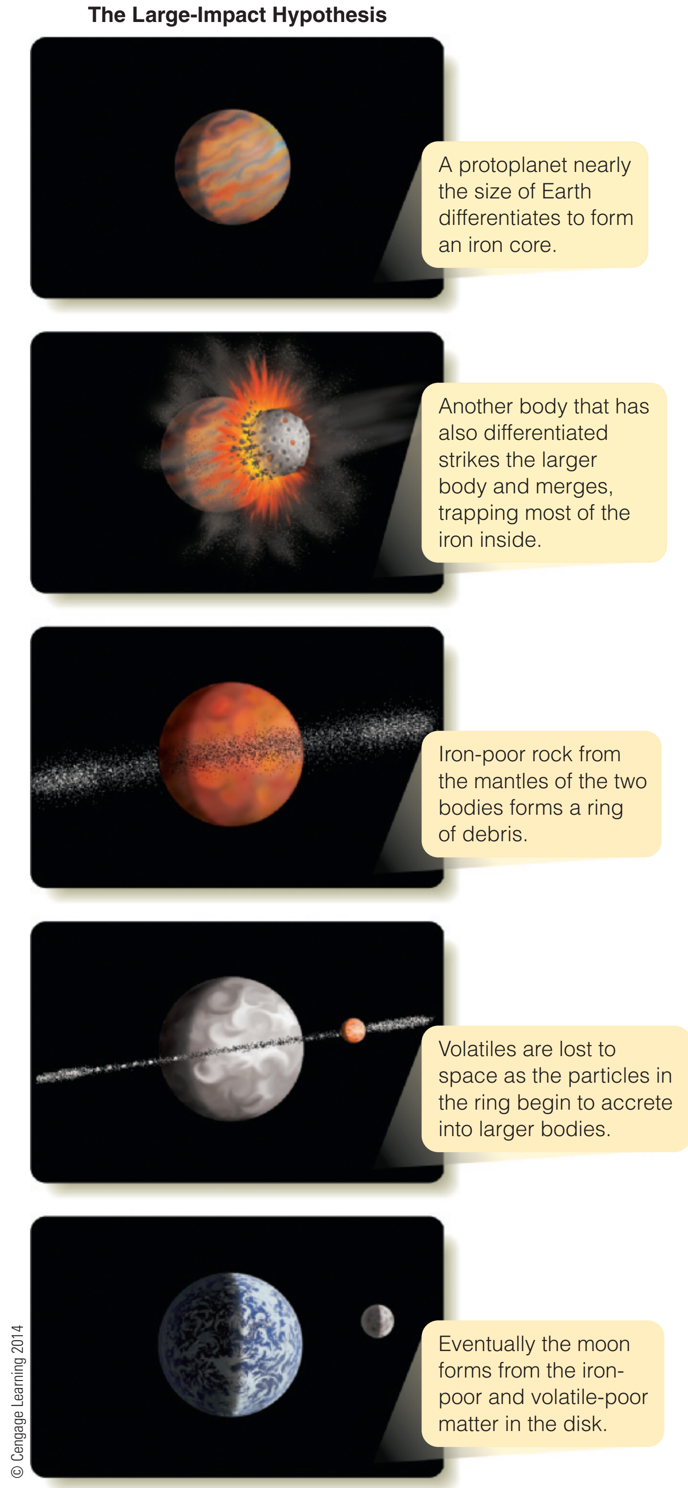
Rocks returned from the moon show that the moon once had a deep magma ocean of molten rock, that after the surface solidified it was heavily fractured by cratering, and that it is now affected mainly by micrometeorites grinding away at surface rock.

The Apollo moon rocks, especially anorthosite from the highlands, show that the moon must have formed in a molten state. Planetary geologists now refer to the exterior of the newborn moon as a **magma ocean.** (*Magma* is the term for molten rock in general, whereas *lava* means molten rock flowing on the surface of a world.) Denser materials sank to the bottom of the magma, and as the magma cooled, low-density minerals floated to the top to form a low-density crust. In this way the moon partly differentiated. The radioactive ages of moon rocks brought back by the Apollo astronauts show that the surface solidified about 4.4 billion years ago.

The second stage, cratering, began as soon as the crust solidified, and the older highlands show that cratering was intense during the heavy bombardment period at the end of planet building. The moon's crust was shattered, and the largest impacts formed giant multiringed crater basins hundreds of kilometers in diameter (■ Figure 9-9). The basin that became Mare Imbrium ("Sea of Rains"), for instance, was blasted out by the impact of an object about the size of Rhode Island. This Imbrium event occurred about 4 billion years ago and blanketed 16 percent of the moon with ejecta.

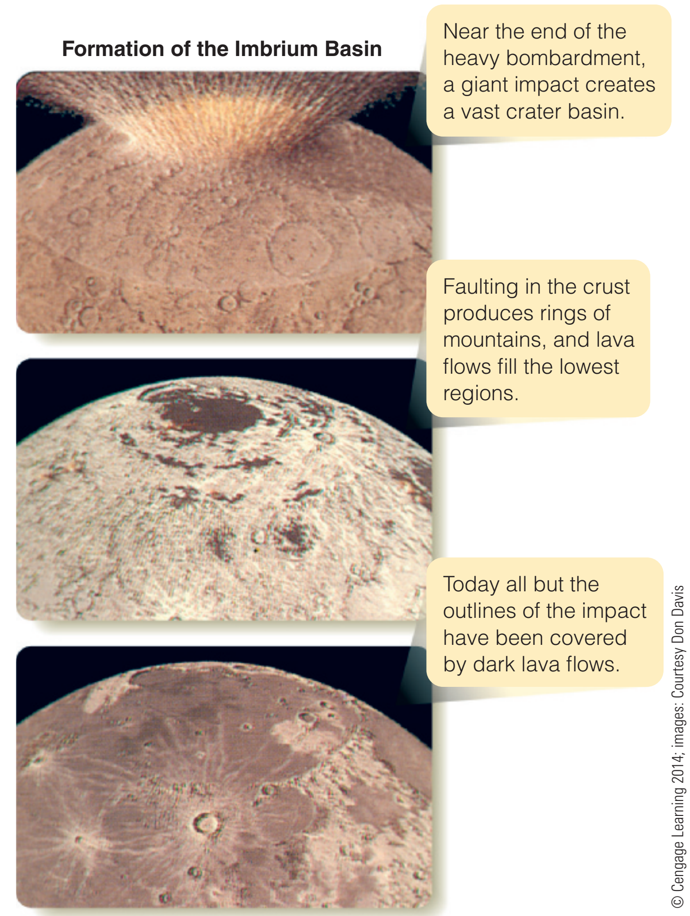
By counting craters on different parts of the moon and calibrating those data using radioactive ages of lunar rock samples, planetary scientists have been able to chart the decline of cratering over the first billion years of the solar system's history. By about 3.8 billion years ago, the impact rate had decreased to the current low rate, but before that there seems to have been a sudden burst of cratering astronomers call the **late heavy bombardment** that might have included the Imbrium event. The primary evidence of this late bombardment comes from observations of the moon, but that sudden storm of impacts must have affected all of the planets in the solar system, including Earth. Those craters on Earth were erased long ago, but they can be seen on other worlds such as Mercury, Mars, and some of the moons in the outer solar system. To understand what astronomers think caused the late heavy bombardment, you will need to learn more about the formation and evolution of Uranus and Neptune in Chapter 11.

Astronomers can calculate that the tremendous impacts that formed the lunar basins would have cracked the crust to depths of 10 kilometers or more and led to the third stage—flooding. Though Earth's moon cooled rapidly after its formation,



■ **Figure 9-8**

When the solar system was about 50 million years old, a massive collision produced the moon in its orbit inclined to Earth's equator.



■ **Figure 9-9**

Mare Imbrium ("Sea of Rains") on the moon has a generally round outline, the consequence of its formation by a giant impact 4 billion years ago.

radioactive decay continued to heat lunar subsurface material, as it did inside Earth. Parts of the lunar mantle and lower crust remelted, producing lava that followed the cracks up into the giant basins (■ Figure 9-10). The basins were flooded by successive lava flows of dark basalts from 3.8 to 3.2 billion years ago, thus forming the maria.

Studies of the moon show that its crust is thinner on the side toward Earth, perhaps due to tidal effects. Consequently, while lava flooded the basins on the Earthward side, it was unable to rise through the thicker crust to flood the lowlands on the far side. One of the largest known impact basins in the solar system is the moon's South Pole—Aitken Basin (Figure 9-10b). It is about 2600 km (1600 mi) in diameter and as deep as 13 km (8 mi) in places, but flooding has never filled it with smooth lava flows to make an obvious mare.

The fourth stage, slow surface evolution, has been more limited on the moon than on Earth because the moon has cooled

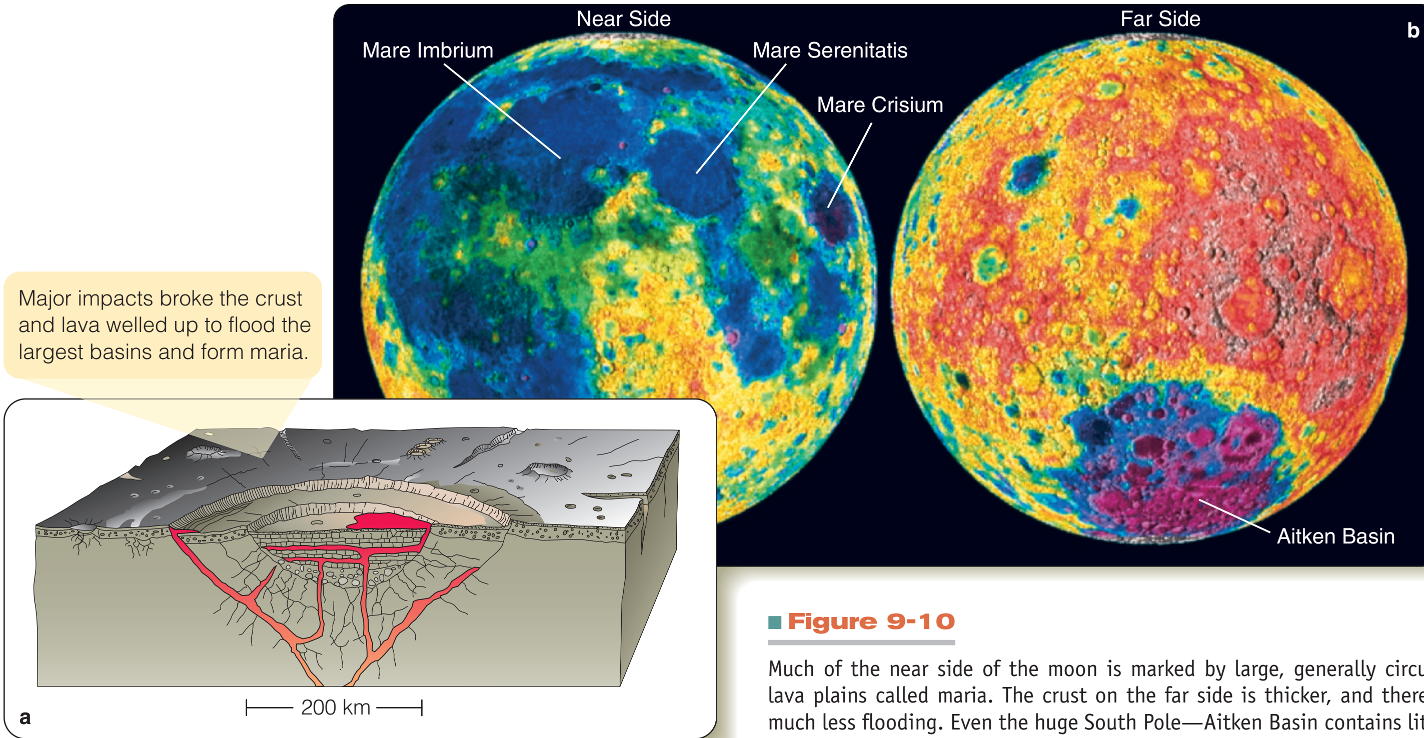


Figure 9-10

Much of the near side of the moon is marked by large, generally circular lava plains called maria. The crust on the far side is thicker, and there is much less flooding. Even the huge South Pole—Aitken Basin contains little lava flooding. In these maps color marks elevation, with red the highest regions and purple the lowest.

© Cengage Learning 2014; Adapted from a diagram by William Hartmann; NASA/Clementine

rapidly and also lacks water. Flooding on Earth included water, but the moon has never had an atmosphere and thus has never had liquid surface water. With no air and no water, erosion is limited to the constant bombardment of micrometeorites and rare larger impacts. As the moon lost its internal heat, volcanism died down, and the moon became geologically dead. Its crust never divided into moving plates—evident from the fact that there are no folded mountain ranges—and the moon is now a one-plate object, frozen between stages 3 and 4.

SCIENTIFIC ARGUMENT

Why are the maria nearly free of craters?

It's been said that timing is everything, and in this case your argument should carefully consider the sequence of events. The evidence from radioactive ages of moon rocks is that the moon's crust formed and was heavily cratered before about 4 billion years ago. The impacts of the heavy bombardment and late heavy bombardment marked the entire lunar surface and created some very large crater basins. Later, after the end of the bombardments, lava welled up and filled the lowlands of the largest crater basins with basalt to form the maria. Craters in the basins were covered over by lava, and there were few impacts afterward to form new craters. Thus, the maria are nearly free of craters, but the ancient highlands remain heavily cratered.

Now build a new argument. **How is timing important in explaining the formation of an iron-poor moon in the large-impact hypothesis?**

What Are We? Scientific Imagineers

One of the most fascinating aspects of science is its power to reveal the unseen. That is, it reveals regions you can never visit. You saw this in earlier chapters when you studied the inside of the sun and stars, the surface of neutron stars, the event horizon around black holes, the cores of active galaxies, and more. In this chapter, you have "seen" Earth's core.

An engineer is a person who builds things, so you can call a person who imagines things an imagineer. Most creatures on Earth cannot imagine situations that do not exist, but humans have evolved the ability to say, "What if?" Our ancient ancestors could imagine what would happen if a tiger was hiding in the grass, and we can imagine the inside of Earth.

A poet can imagine the heart of Earth, and a great writer can imagine a journey to the center of Earth. In contrast, scientists use their imagination in a carefully controlled way. Guided by evidence and theory, they can imagine the molten core of our planet. As you read this chapter, if you saw the yellow-orange glow and felt the heat of the liquid iron, then you were a scientific imagineer.

Human imagination makes science possible and provides one of the great thrills of science—exploring beyond the limits of normal human experience.

Study and Review

Summary

- Earth is the standard of **comparative planetology (p. 166)** in the study of the Terrestrial planets because we know it best and because it contains all of the phenomena found on the other Terrestrial planets.
- Our discussion of the Terrestrial planets considers Earth, the moon, Mercury, Venus, and Mars. Earth's moon is included because it is a complex world and makes a striking comparison with Earth.
- The Terrestrial worlds differ mainly in size, but they all have low-density crusts, **mantles (p. 166)** of dense rock, and metallic cores.
- Comparative planetology warns you to expect that cratered surfaces are old, that heat flowing out of a planet drives geological activity, and that the nature of a planet's atmosphere depends on the size of the planet and its temperature.
- At some point early in its history Earth was hot enough to be completely molten, which caused it to differentiate into layers of different density.
- Earth has passed through four stages as it evolved: (1) differentiation, (2) cratering, (3) flooding by lava and water, and (4) slow surface evolution. The other Terrestrial planets and the moon also passed through the same stages, which had different effects and durations depending on the specific properties of each body.
- Earth is unique in the solar system in that it has large amounts of liquid water on its surface, and that water drives strong erosion that alters the surface geology. Earth is also unique in that it is the only known home for life.
- **Seismic waves (p. 169)** generated by earthquakes can be detected by **seismographs (p. 169)** all over the world and can reveal Earth's internal structure.
- **Pressure *(P)* waves (p. 169)** can travel through a liquid, but **shear *(S)* waves (p. 170)** cannot. Observations show that *S* waves cannot pass through Earth's core, and that is evidence that the core is liquid. Measurements of heat flowing outward from the interior, combined with mathematical models, reveal that the core is very hot and composed of iron and nickel.
- Although Earth's crust is brittle and breaks under stress, the mantle is **plastic (p. 170)** and can deform and flow under pressure.
- Earth is dominated by **plate tectonics (p. 172),** which breaks the crust into moving sections. Plate tectonics is driven by heat flowing upward from the interior.
- Tectonic plates are made of low-density, brittle rock that floats on the hotter plastic upper layers of the mantle. **Rift valleys (p. 172)** can be produced where plates begin pulling away from each other.
- New crust is formed along **mid-ocean rises (p. 172),** where molten rock solidifies to form **basalt (p. 172).** Crust is destroyed when it sinks into the mantle along **subduction zones (p. 172).** Volcanism and earthquakes are common along the edges of the plates.
- The motion of a plate across a hot spot can produce a chain of volcanic islands such as the Hawaiian Island chain. Hot-spot volcanism is not related to subduction zones.
- The continents are drifting slowly on the plastic mantle, and their arrangement changes with time. Where they collide, they can form **folded mountain ranges (p. 172).**
- Most geological features on Earth, such as mountain ranges and the Grand Canyon, have been formed recently. The first billion years of Earth's geology are almost entirely erased by plate tectonics and erosion.
- Because Earth formed in a molten state, its **primary atmosphere (p. 171)** was probably mostly carbon dioxide, nitrogen, and water vapor. Most of the carbon dioxide eventually dissolved in seawater and was added to ocean sediments, and plant life has added oxygen to the atmosphere, producing the present **secondary atmosphere (p. 171).**
- Ultraviolet photons can break up water molecules in a planet's atmosphere, but as soon as Earth had enough oxygen, an **ozone layer (p. 171)** could form high in Earth's atmosphere. The ozone absorbs ultraviolet photons and protects water molecules.
- The **albedo (p. 174)** of a planet is the fraction of sunlight hitting it that it reflects into space. Small changes in the albedo of Earth caused by changes in clouds and atmospheric currents can have a dramatic effect on climate.
- The **greenhouse effect (p. 171)** can warm a planet if gases such as carbon dioxide in the atmosphere are transparent to light but opaque to infrared. The natural greenhouse effect warms Earth and makes it comfortable for life, but greenhouse gases added by industrial civilization are responsible for **global warming (p. 171).**
- Measurement of carbon isotope ratios and carbon dioxide versus oxygen abundances make it clear that the CO_2 added to the atmosphere since 1800 is predominantly from burning of fossil fuels. Observations and model calculations have eliminated other candidate causes for the current warming such as natural climate cycles or variations in the sun's output. The precise amount of future warming is difficult to predict.
- The ozone layer high in Earth's atmosphere protects the surface from ultraviolet radiation, but certain chemicals called chlorofluorocarbons released in industrial processes attack the ozone layer and thin it. This is allowing more harmful ultraviolet radiation to reach Earth's surface.
- Earth's moon is a small world and so has lost most of its internal heat and is no longer geologically active. It contains little metal and has an overall low density.
- The overall albedo of the moon is very low, even in the relatively bright highlands.
- The moon's surface is fractured by impacts, producing craters that are especially easy to see near the **terminator (p. 177),** the moving boundary between the sunlit and unlit parts of the moon.
- Debris blasted out of craters is called **ejecta (p. 178)** and can produce **rays (p. 178)** and **secondary craters (p. 178).** Large impacts have resulted in prominent **multiringed basins (p. 179).** Constant bombardment by tiny **micrometeorites (p. 179)** continues to erode the moon, producing a layer of dust on the surface.
- The moon's highlands are the oldest portions of its surface, heavily cratered by impacts that occurred during the heavy bombardment of all the planets after the formation of the solar system. There is evidence from crater counts and lunar rock sample ages that, before the bombardment ended, there was a temporary burst of impacts called the **late heavy bombardment (p. 181).**
- The moon's lowlands are filled by lava flows that formed smooth **maria** (singular, **mare**) **(p. 177)** soon after the end of the heavy bombardments.
- Lunar rocks brought back to Earth include **vesicular basalts (p. 177)** from parts of the moon where the surface was covered by successive lava flows; **anorthosite (p. 177),** which shows that a large portion of the moon was once a **magma ocean (p. 181);** and **breccias (p. 177),** providing evidence that the moon's crust has been repeatedly pounded by impacts.
- The **large-impact hypothesis (p. 180)** suggests the moon formed when an impact between the proto-Earth and a very large planetesimal left Earth surrounded with a disk of collision debris. The moon formed from that disk.

Study and Review

Review Questions

1. Why would you include the moon in a comparison of the Terrestrial planets?
2. In what ways is Earth unique among the Terrestrial planets?
3. What are the four stages of planetary development?
4. How do you know that Earth differentiated?
5. How are earthquakes in Hawai'i different from those in Southern California?
6. What characteristics must Earth's core have in order to generate a magnetic field?
7. How do island chains located in the centers of tectonic plates such as the Hawaiian Islands help you understand plate tectonics?
8. What has produced the oxygen in Earth's atmosphere?
9. How does the increasing abundance of CO_2 in Earth's atmosphere cause a rise in Earth's temperature?
10. Why would a decrease in the density of the ozone layer cause public health problems?
11. Why doesn't Earth have as many craters as the moon?
12. What kind of erosion is now active on Earth's moon?
13. Discuss the evidence and hypotheses concerning the origin of Earth's moon.
14. **How Do We Know?** Why is heat flow the key to understanding a planet's surface activity?
15. **How Do We Know?** In what ways have scientists been professionally skeptical about global warming and its causes?

Discussion Questions

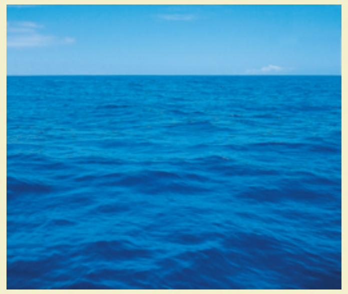
1. If you orbited a planet in another solar system and discovered oxygen in its atmosphere, what might you expect to find on its surface?
2. If liquid water is rare on the surface of planets, then most Terrestrial planets must have CO_2-rich atmospheres. Why?
3. Old science-fiction paintings and drawings of colonies on the moon often showed very steep, jagged mountains. Why did the artists assume that lunar mountains would be more jagged than mountains on Earth? Why are lunar mountains actually less jagged than mountains on Earth?

Problems

1. Assume *P* waves travel at 10 km/s and *S* waves travel at 5 km/s. If the *S* waves from an earthquake arrive at a seismographic station 10 minutes after the *P* waves, how far away was the earthquake from the station?
2. What percentage of Earth's volume is taken up by its metallic core?
3. If the Atlantic seafloor is spreading at 3.0 cm/year and is now 6400 km wide, how long ago were the continents in contact?
4. The Hawaiian-Emperor chain of undersea volcanoes is about 7500 km long, and the Pacific plate is moving 9.2 cm a year. How old is the oldest detectable volcano in the chain? What has happened to older volcanoes in the chain?
5. Calculate the age of the Grand Canyon as a percent of Earth's age.
6. Earth is four times larger in diameter than its moon. How many times larger is it in surface area? In volume?
7. The smallest detail visible through Earth-based telescopes is about 1 arc second in diameter. What size object would this represent on Earth's moon? (*Hint:* Use the small angle formula, Chapter 3.)

Learning to Look

1. Look at the globe of Earth shown on page 173 and look for volcanoes scattered over the Pacific Ocean. What is producing these volcanoes?
2. In what ways is the photo at the right a typical view of the surface of planet Earth? How is it unusual among planets in general?

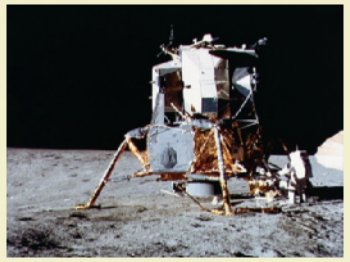


William K. Hartmann/Planetary Science Institute

3. What do you see in this photo that suggests heat is flowing out of Earth's interior?

USGS

4. In the photo to the right, Astronaut Alan Bean works at the Apollo 12 lander *Intrepid*. Describe the surface you see. What kind of terrain did they land on, for this, the second human landing on the moon?



NASA

5. Examine the mountains at the Apollo 17 landing site (Figure 9-6). What processes shape mountains on Earth that have not affected mountains on the moon?

Great Debates

1. *Heritage Sites on the Moon.* World Heritage Sites are places on Earth that have special cultural or physical significance. Since commercial spacecraft are being built to go to the Moon, should places on the Moon (for example, first footprints, first landing site, and the like) be listed as Heritage Sites? What penalties should be issued, and by whom, if a person fails to abide by any Heritage Site rules? Who (which country or organization) should fly to the moon to defend the sites if they are deemed heritage-worthy, and who pays for the flight?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
2. *Space Travel and Contamination.* Suppose one group of tourists or astronauts visits the moon. Suppose one or more of the visitors take a memento that causes the inhabitants to become sick on the return trip. They radio ahead to inform Earth that the sick need medical attention. A group, which includes you, convenes to decide whether to allow the spacecraft to return to Earth. No one knows if the toxins on the spacecraft could eliminate the human race. Should the risk be taken to allow the spacecraft to land?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find information on NASA's efforts to protect life on Earth.
   c. Cite your sources.
3. *Earth's Rotation Axis Shift.* In popular literature, the sudden geographic movement of Earth's north and south poles results in a new tilt angle of Earth's rotation axis, causing disasters such as floods and quakes, for example. Although these works are considered fiction, should the authors be concerned that their readers may actually believe their ideas? Should popular fiction writers take on more of an educational role to their readers? Would a work of fiction be made better if the work were more accurate in the science it portrays?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
4. *Geomagnetic Reversal.* Earth's magnetic poles have changed polarity in the past; the positions of magnetic north and south change positions several times per million years. The most recent magnetic pole switch happened 780,000 years ago, meaning Earth may be overdue for a magnetic reversal. Some have hypothesized that mass extinctions may occur as a result of a pole reversal due to excessive volcanism and outgassing that could blanket the Earth, smothering surface life. Should governments like the United States have a disaster plan in place should such effects from a pole reversal be possible? What would be some of the steps that a government might take to save lives in a situation like this? Is saving an entire country too much for one government to handle?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
5. *Plate Tectonics.* Are plate tectonics necessary for life and especially for intelligent life on Earth? Should taxpayer dollars be spent on studying or visiting celestial objects that do not have plate tectonics?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

ENHANCED WebAssign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Seismic Waves
- The Moon's Craters
- The Origin of the Moon

CengageNOW Virtual Astronomy Labs 2.0

Virtual ASTRONOMY LABORATORIES

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 5: Planetary Geology

From before recorded history to the present day, humans have watched in awe and fear as volcanoes erupt red-hot lava and geysers spray superheated steam driven by some hellish heat source underground. Miners working in deep tunnels testify that the temperature of the rock increases the farther below the surface they go. But the deepest mines or holes of any sort ever drilled have penetrated barely more than 1/1000 of the way to the center of the planet. What is in Earth's interior, farther down than anyone has been able to explore in person?

As you learned in an earlier chapter, temperature corresponds to the average speed of motion of atoms in a material. Thermal (heat) energy always flows from hot places to cooler places. In human affairs, "follow the money" is a slogan referring to the fact that, unless you know who is paying and who is benefiting from a financial transaction, you don't really know what is going on. In geologic affairs, the equivalent slogan is "follow the energy." The variety of ways in which the heat energy of Earth and other planets travels from high temperature interiors to cooler surfaces tell you what's really driving the processes and changes you can observe.

Density is another crucial measure of what is inside Earth and other planets. Density is defined as mass per volume. A brick of concrete and a brick of foam padding may have exactly the same size, but the concrete brick has more mass in that volume than does the foam and thus a higher density. Knowing the density of an unknown material limits possible choices for its composition; the exact density of the concrete brick tells you it must be concrete or other similar material. The densities of Earth's surface rocks are typically about 3 or 3.5 grams per cubic centimeter. But the overall density of Earth, determined by dividing the planet's total mass by its total volume, is 5.5 grams per cubic centimeter. So, simple arithmetic tells you that whatever is inside the core of Earth must be much denser, with a different composition, than surface rocks.

In this chapter you found that scientists have developed fairly detailed models of Earth's interior. Those models are based on the principle, verifiable in laboratory experiments, that the speed of seismic waves depends on the temperature and density of the material through which they pass. Seismic waves also change their direction of motion when they run obliquely into boundaries between zones of different density. And, as you learned from this chapter, one type of seismic wave, *S* ("shear") waves, is completely blocked by liquid. After many decades of studying seismic waves traveling from earthquakes to recording stations along thousands of chords drawn across the inside of the globe, scientists are sure that Earth's outer core is liquid but that the mantle above it and the inner core below it are both solid or semisolid. Earth's interior conditions have been plumbed remotely almost as well as if we had traveled there.

Section 1 of Virtual Astronomy Lab 5, "Planetary Geology," reminds you about how density is calculated, then shows you how observations of seismic waves have revealed layers of different density and composition inside Earth. Section 2 reviews conditions and processes on Earth's surface that confirm models of the interior. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 4: Solar Wind and Cosmic Rays

Magnets are fascinating to children and some adults. Albert Einstein attributed the beginning of his interest in the natural world to his amazement at seeing a magnet reach out through space to attract and hold pieces of metal—but only metal, and only certain types of metals. Perhaps your imagination, along with Einstein's, can envision a "field" that is a region filled with the magnet's invisible influence.

Humans have known Earth has a magnetic field (although they didn't call it that) since ancient times. A magnetic needle points approximately north–south and so can be used to make a compass. Some people have imagined that this is due to a huge lodestone (magnetite) mountain near the North Pole. But then careful experiments by physicists over many decades showed that magnetism and electricity are closely related: Electrical currents create magnetic fields. So what electrical current is producing Earth's magnetic field? You learned in this chapter that the existence of Earth's magnetic field is one of several pieces of evidence pointing to the presence of a liquid metal component in Earth's core. Apparently, movement of that electrically conducting fluid is the current that makes the planet's magnetic field.

In an earlier chapter about the sun you learned to imagine interplanetary space as being full rather than empty. Earth and the other planets are embedded in the solar wind, a flow of ionized atoms escaping from the extremely hot outer part of the sun's atmosphere, the corona. You also learned that magnetic fields influence the motions of charged particles. When the electrically charged solar wind encounters Earth's magnetic field, some of the wind particles are forced to go around Earth's magnetosphere—the region of influence of Earth's magnetic field—like water in a stream goes around a rock. Other solar wind particles become trapped in Earth's field in regions named the Van Allen belts after their discoverer, physicist James Van Allen. And some particles move along a path of least resistance, zooming down toward Earth's magnetic poles and colliding with atoms in Earth's upper atmosphere, causing the spectral emission line glows of the aurora borealis and aurora australis (Northern and Southern Lights).

Section 2 of Virtual Astronomy Lab 4, "Solar Wind and Cosmic Rays," guides you to understanding about how magnetic fields affect the motions of charged particles, and how the solar wind interacts with Earth's magnetic field. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

10

Mercury, Venus, and Mars

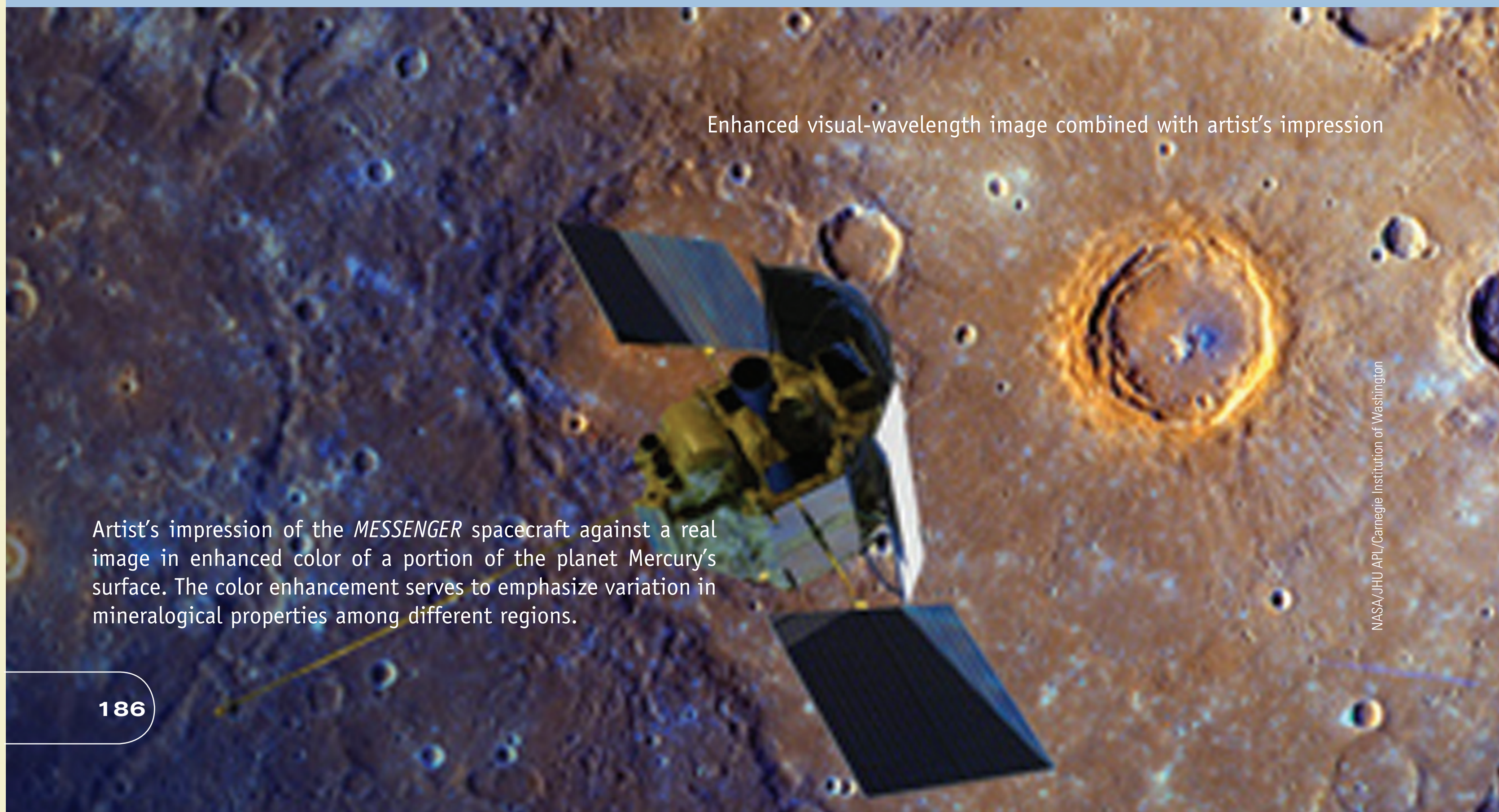
Guidepost

In the previous chapter, you compared a geologically active planet, Earth, the largest of the Terrestrial worlds, with the smallest, Earth's geologically inactive moon. In this chapter, as you continue studying individual Terrestrial planets, you can continue comparing the planets with each other, searching for similarities and contrasts. Like people, the Terrestrial planets are more alike than they are different, but it is the differences that are most memorable.

As you explore, you will be searching for answers to four important questions:

- **How is Mercury similar to, and different from, Earth's moon?**
- **How does distance from the sun affect a planet and its atmosphere?**
- **How does size determine the geologic activity and evolution of a planet?**
- **What is the evidence that Venus and Mars were once more Earth-like, and why did they change?**

Once you have finished exploring the Terrestrial planets, you will be ready to meet a stranger group of characters in the next chapter, the worlds of the outer solar system.



Artist's impression of the *MESSENGER* spacecraft against a real image in enhanced color of a portion of the planet Mercury's surface. The color enhancement serves to emphasize variation in mineralogical properties among different regions.

The only truly alien planet is Earth.

J. G. BALLARD

10-1 Mercury

The moon's Celestial Profile from the previous chapter is repeated here next to **Celestial Profile 4** for Mercury (page 188) to help you do comparative planetology, finding the similarities and differences between the two objects.

Spacecraft Visiting Mercury

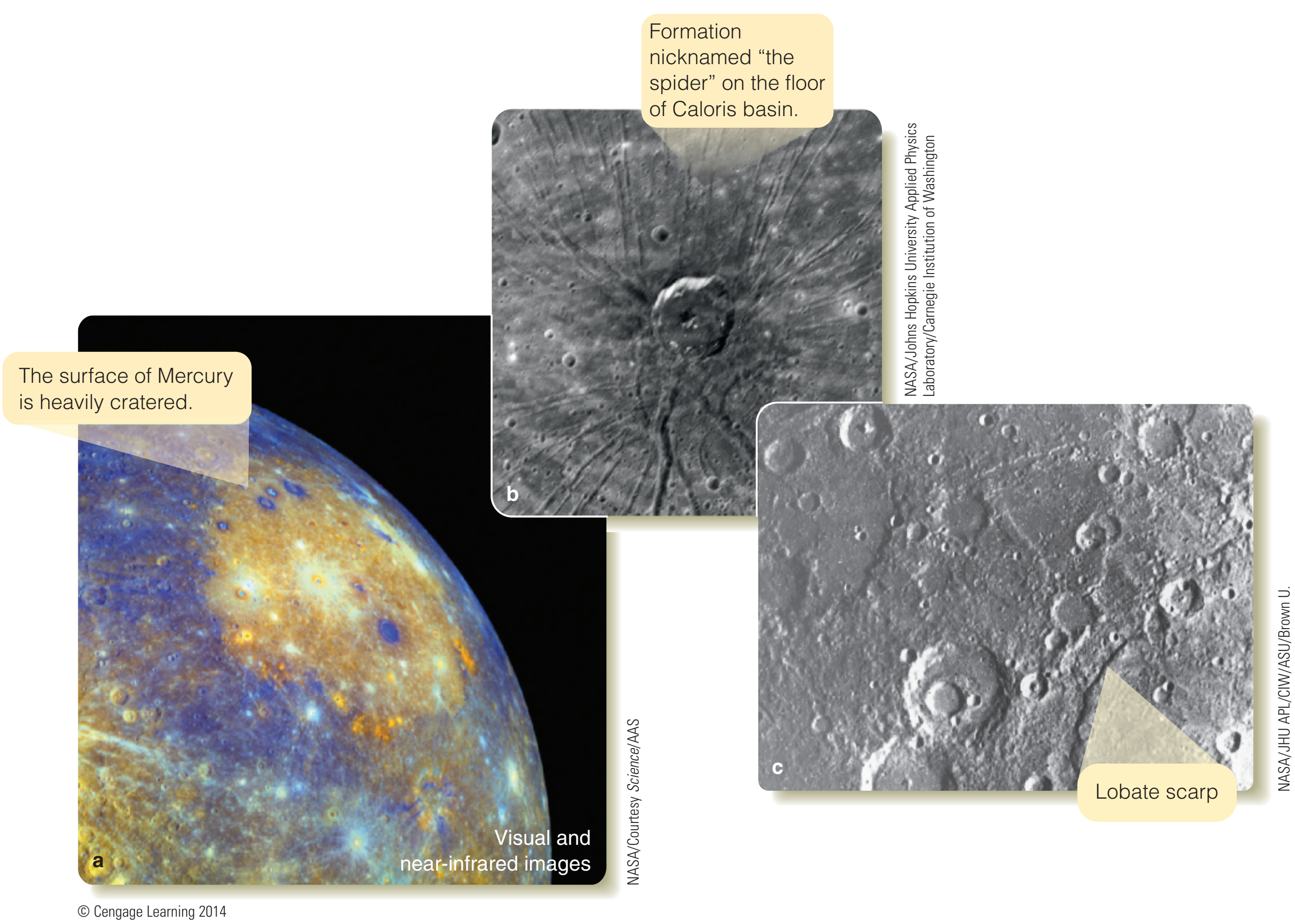
Mercury orbits so close to the sun that it is difficult to observe from Earth, and little was known about it until 1974–1975, when the *Mariner 10* spacecraft flew past Mercury three times and revealed a planet with a heavily cratered surface resembling a larger version of Earth's moon. Analysis of the *Mariner 10* data showed that large areas have been flooded by lava and then cratered. Spectacular new images and data arrived from the *MESSENGER* spacecraft that flew by Mercury three times before finally settling into orbit around the planet in 2011 (■ Figure 10-1).

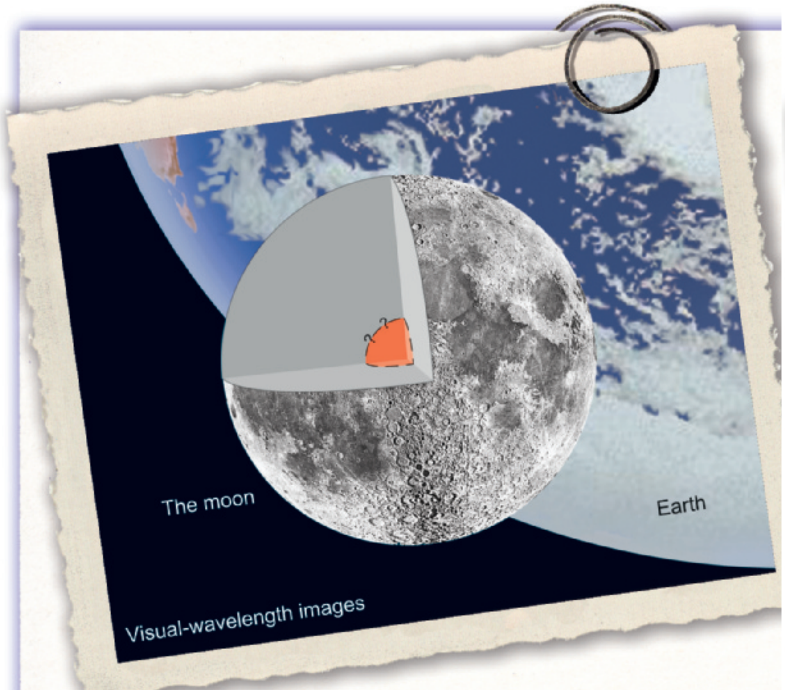
The largest impact feature on Mercury is the Caloris Basin, a mountain-ringed basin that *MESSENGER* photos reveal as over 1500 km (900 mi) in diameter (Figure 10-1a), resembling the large ringed basin Mare Orientale ("Eastern Sea") on Earth's moon. The Caloris Basin on Mercury and Mare Orientale on the moon both include concentric rings of cliffs formed by a large impact.

Though Mercury looks moonlike, it does have several features that Earth's moon lacks. *MESSENGER* photos revealed a "spider" (Figure 10-1c) of raised ridges appearing to extend from near a medium-sized crater; geologists are not sure what process could have caused the spider. Images from both *Mariner 10* and

■ Figure 10-1

(a) Enhanced-color mosaic of images from the *MESSENGER* spacecraft of the portion of the planet Mercury containing the Caloris multiringed basin. The color enhancement emphasizes variations in composition among different parts of the surface. (b) The origin of the "spider" formation photographed by *MESSENGER* is a puzzle. (c) Lobate scarps cross craters, indicating that Mercury cooled and shrank, wrinkling its crust, after many of the craters had formed.





© Cengage Learning 2014; image: NASA

Earth's moon has about a quarter the diameter of Earth. Its low density indicates that it does not contain much iron. The size of its core, if any, and the amount of remaining heat are unknown.

Celestial Profile 3: The Moon

Motion:

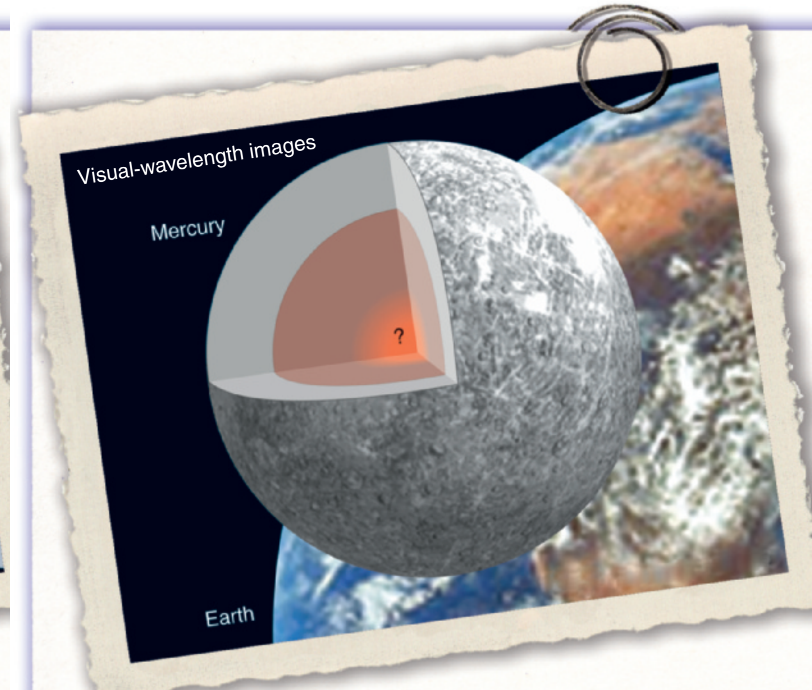
Average distance from Earth	3.84×10^5 km (center to center)
Eccentricity of orbit	0.055
Inclination of orbit to ecliptic	5.1°
Orbital period (sidereal)	27.3 d
Synodic period (phase cycle)	29.5 d
Inclination of equator to orbit	6.7°

Characteristics:

Equatorial diameter	3.48×10^3 km (0.273 $D_{\oplus}$)
Mass	7.35×10^{22} kg (0.0123 $M_{\oplus}$)
Average density	3.35 g/cm^3 (3.3 g/cm^3 uncompressed)
Surface gravity	0.17 Earth gravity
Escape velocity	2.4 km/s (0.21 $V_{\oplus}$)
Surface temperature	−170° to 130°C (−275° to 265°F)
Average albedo	0.12
Oblateness	0

Personality Point:

Lunar superstitions are very common. *Lunatic* and *lunacy* come from *luna*, the moon. Someone who is *moonstruck* is supposed to be at least a bit nutty. Because the moon affects the ocean tides, many superstitions link the moon to water, to weather, and to women's cycle of fertility. Moonlight is supposed to be harmful to unborn children, but on the plus side, moonlight rituals supposedly can remove warts.



© Cengage Learning 2014; image: NASA

Mercury has slightly over a third the diameter of Earth. Its high density means it must have a very large iron core. The amount of heat that Mercury retains is unknown.

Celestial Profile 4: Mercury

Motion:

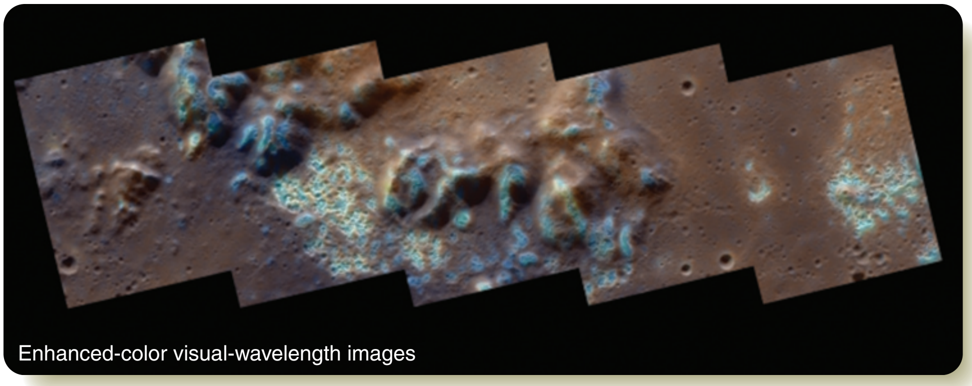
Average distance from the sun	0.387 AU (5.79×10^7 km)
Eccentricity of orbit	0.206
Inclination of orbit to ecliptic	7.0°
Orbital period	0.241 y (88.0 d)
Period of rotation (sidereal)	58.6 d
Inclination of equator to orbit	0.0°

Characteristics:

Equatorial diameter	4.89×10^3 km (0.382 $D_{\oplus}$)
Mass	3.30×10^{23} kg (0.0553 $M_{\oplus}$)
Average density	5.43 g/cm^3 (5.4 g/cm^3 uncompressed)
Surface gravity	0.38 Earth gravity
Escape velocity	4.3 km/s (0.38 $V_{\oplus}$)
Surface temperature	−170° to 430°C (−275° to 805°F)
Average albedo	0.12
Oblateness	0

Personality Point:

Mercury lies very close to the sun and completes an orbit in only 88 Earth days. For this reason, the ancients named the planet after Mercury, the fleet-footed messenger of the gods. The name is also applied to the element mercury, which is known as quicksilver because it is a heavy, quick-flowing silvery liquid at room temperatures.



■ **Figure 10-2**

Mosaic of *MESSENGER* enhanced-color images of a section of the floor and peak-ring mountains of the Raditladi impact basin on Mercury. The individual frames in the mosaic are about 20 km (12 mi) wide. The rounded depressions called "hollows," seen in many locations on Mercury, may have been formed by sublimation of a volatile component in the surface material.

MESSENGER reveal long curving ridges called **lobate scarps** up to 3 km (2 mi) high and 500 km (300 mi) long (■ Figure 10-1c). The scarps even cut through craters, indicating that they formed after most of the heavy bombardment. The lobate scarps are the kind of faults that form by compression. This suggests that the entire crust of Mercury was compressed long ago.

Mercury is quite dense, and models indicate that it must have a large metallic core. In fact, the metallic core occupies about 70 percent of the radius of the planet. In a sense, Mercury is a metal planet with a thin rock mantle and crust. Spectroscopic observations indicate that Mercury has an extremely thin atmosphere that may be partly outgassed from the crust and partly captured from the solar wind.

The History of Mercury

The accumulated facts about Mercury don't really help you understand the planet until you have a unifying hypothesis. Like a story, it must make sense and bring the known facts together in a logical argument that explains how Mercury got to be the way it is (**How Do We Know? 10-1**).

Mercury is small, and that fact has determined much of its history. Like Earth's moon, Mercury has lost much of its internal heat, and thus is no longer geologically active.

In the first stage of its planetary history, Mercury differentiated to form a metallic core and a rocky mantle. The presence of a magnetic field about 10^{4} as strong as Earth's is further evidence of a metallic core. You learned in Chapter 8 that the condensation sequence could explain a high abundance of metals in Mercury, but detailed calculations show that Mercury contains even more iron than the condensation sequence would predict. Drawing on the large-impact hypothesis for the origin of Earth's moon, scientists proposed that Mercury suffered a major impact after differentiation that drove away much of the rocky mantle. However, recent data from *MESSENGER*, including images of surface hollows in many locations (■ Figure 10-2), indicate that a substantial amount of volatiles were incorporated in the crust and subsequently vaporized. Abundant volatiles would not have survived a giant impact, so an impact may not, after all, be the explanation for the size of Mercury's metal core.

In the second and third stages of planet formation, cratering battered Mercury's crust, and lava flows welled up to fill the lowlands, just as they did on the moon. As Mercury lost internal heat, its large metal core contracted and its crust was compressed, breaking to form the lobate scarps, much as the peel of a drying apple wrinkles.

Mercury is now a one-plate planet much like Earth's moon and, lacking a significant atmosphere to erode its surface, has changed little since the last lava hardened.

SCIENTIFIC ARGUMENT

Why does Mercury have lobate scarps, but Earth does not?
At first glance, you might build an argument to propose that any world with a metallic interior should have lobate scarps, but other factors are also important. Earth has a fairly large metallic core, but being a large world, it has not cooled very much, so it presumably hasn't shrunk much. Also, the geologic activity on Earth's surface would have erased such scarps if they formed long ago. On the other hand, Earth's moon is not geologically active and does not contain a large metallic core, but faint lobate scarps have been found there. Evidently, the moon shrank a bit as it lost its internal heat.

Now expand your argument. **How do you know the lobate scarps formed after most of the heavy bombardment was over?**

10-2 Venus

You might expect Venus to be much like Earth. Its diameter is 95 percent of Earth's (**Celestial Profile 5**, page 197), it has a similar average density and composition, and it is just 30 percent closer to the sun. Unfortunately, the surface of Venus is perpetually hidden below thick clouds, and only in the past few decades have planetary scientists discovered that Venus is a deadly hot desert world of volcanoes, lava flows, and impact craters lying within

How Do We Know? | 10-1

Hypotheses and Theories Unify the Details

How do scientists make sense out of all the details? Like any technical subject, science includes a mass of details, facts, figures, measurements, and observations. It is easy to be overwhelmed by the flood of details, but one of the most important characteristics of science comes to your rescue. The goal of science is not to discover more details but to explain the details with a unifying hypothesis or theory. A good theory is like a basket that makes it easier for you to carry a large assortment of details.

This is true of all the sciences. When a psychologist begins studying the way the human eye and brain respond to moving points of light, the data are a sea of detailed measurements and observations. Once the psychologist forms a hypothesis about the way the eye and brain interact, the details fall into place as parts of a logical story. If you understand the hypothesis, the details all fit together and make sense, and thus you can remember the details without blindly memorizing tables of facts and figures. The goal of science is understanding, not memorization.

Scientists are in the storytelling business. The stories are often called hypotheses or theories, but they are, in a sense, just stories to explain how nature works. The difference between scientific stories and works of fiction lies in the use of facts. Scientific stories are constructed to fit all known facts and are then tested over and over against new facts obtained by observation and experiment.

When you try to tell the story of each planet in our solar system, you pull together all the hypotheses and theories and try to make them into a logical history of how the planet got to be the way it is. Of course, your stories will be incomplete because scientists don't understand all the factors in planetary evolution. Nevertheless, your story of each planet will draw together the known facts and details and attempt to make them into a logical whole.

Memorizing a list of facts can give you a false feeling of security, just as when you memorize the names of things without understanding them. Rather than memorizing facts, you should search for the unifying hypothesis that pulls the details together into a single story. Your goal in studying science should be to understand nature, not just to remember facts.

Phyllis Leber

When scientists create a hypothesis, it draws together a great many observations and measurements.

a deep atmosphere of hot gases. No spacesuit is tough enough to allow you to visit the surface of Venus.

The Atmosphere of Venus

In composition, temperature, and density, the atmosphere of Venus is more Hades than Heaven. The air is unbreathable, very hot, and almost 100 times denser than Earth's air. How do planetary scientists know this? Because U.S. and Soviet space probes descended into the atmosphere and, in a few cases, landed and reported back from the surface.

In composition, the atmosphere of Venus is roughly 96 percent carbon dioxide. The rest is mostly nitrogen, with some argon, sulfur dioxide, and small amounts of sulfuric acid, hydrochloric acid, and hydrofluoric acid. There is only a tiny amount of water vapor. On the whole, the composition is deadly unpleasant, and most certainly smells bad too. Spectra show that the impenetrable clouds that hide the surface are made up of droplets of sulfuric acid and microscopic crystals of sulfur (■ Figure 10-3).

This strange atmosphere is 90 times denser than Earth's atmosphere. The air you breathe is 1000 times less dense than water, but on Venus the air is only 10 times less dense than water. If you could survive the unpleasant conditions, you could strap wings on your arms and fly in Venus's atmosphere.

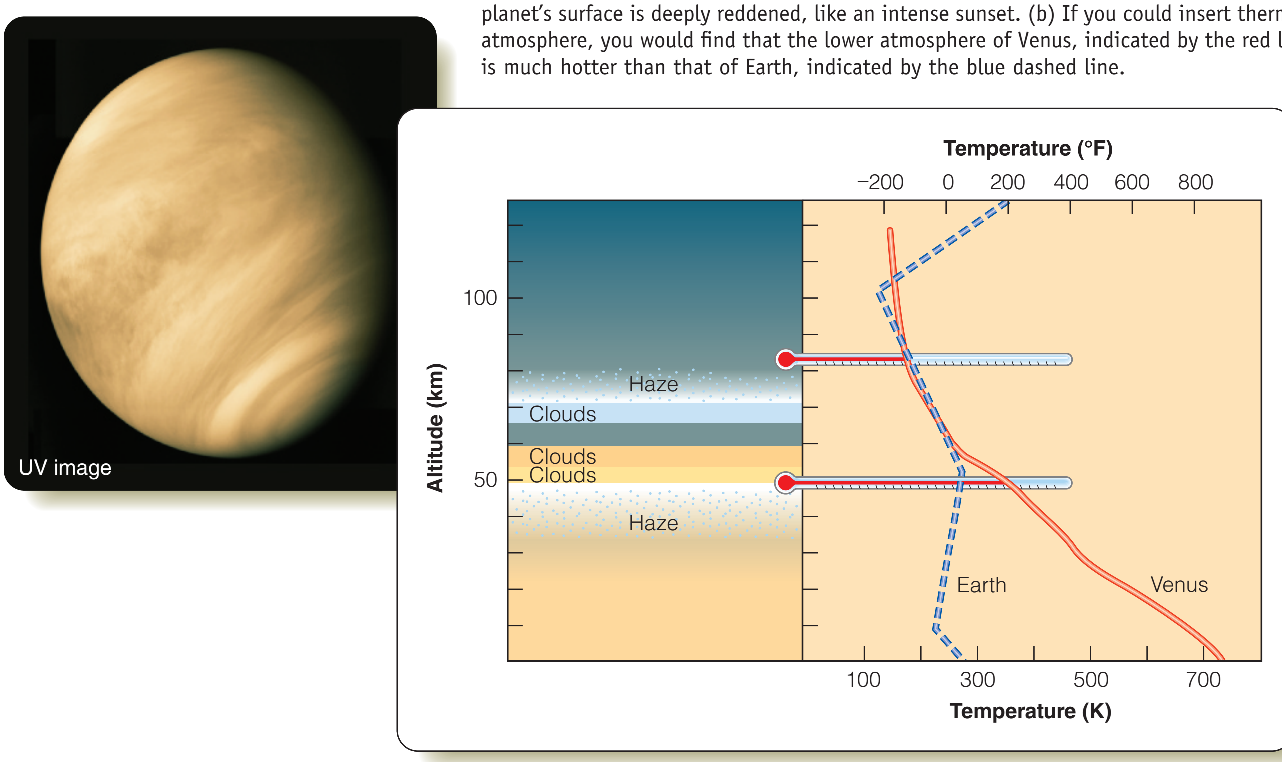
The surface temperature on Venus (look again at Celestial Profile 5) is hot enough to melt lead, and you can understand that because the thick atmosphere creates a severe greenhouse effect. Sunlight filters down through the clouds and warms the surface, but heat cannot escape easily because the atmosphere is opaque to infrared radiation. Traces of sulfur dioxide and water vapor help trap the infrared, but it is the overwhelming abundance of carbon dioxide that makes the greenhouse effect on Venus much more severe than on Earth.

The Surface of Venus

Although the thick clouds on Venus are opaque to visible light, they are transparent to radio waves, so astronomers have been able to map Venus using radar. As early as 1965, Earth-based radio telescopes made low-resolution maps, but later both U.S. and Soviet spacecraft orbited Venus and mapped its surface by radar. Maps made in the early 1990s by the *Magellan* spacecraft reveal objects as small as 100 meters (300 feet) in diameter.

■ **Figure 10-3**

(a) The three main cloud layers in the atmosphere of Venus are more than 10 times higher above the surface than are Earth clouds and completely hide the surface. Venus's clouds are composed mostly of sulfuric acid droplets and reflect much of the sunlight back into space. The light that reaches the planet's surface is deeply reddened, like an intense sunset. (b) If you could insert thermometers into the atmosphere, you would find that the lower atmosphere of Venus, indicated by the red line on the graph, is much hotter than that of Earth, indicated by the blue dashed line.



NASA

© Cengage Learning 2014

Radar maps of Venus are reproduced using arbitrary colors. In some maps, scientists have chosen to give Venus an overall orange glow because sunlight filtering down through the clouds bathes the landscape in a perpetual sunset glow. Other radar maps have been colored gray, the natural color of the rocks. In yet other maps, lowlands are colored blue, but there are no oceans on Venus. When you look carefully at colored radar maps of Venus, recall that its surface is a deadly dry desert.

By international agreement, names on Venus are all female, with three exceptions—Maxwell, a high mountain, and Alpha Regio and Beta Regio, two high volcanic peaks—which were all named before the international naming convention for Venus was adopted.

Radar maps show that Venus is similar to Earth in one way but strangely different in other ways. Nearly 75 percent of Earth's surface is covered by low-lying, basaltic seafloors, and 85 percent of Venus's surface is covered by basaltic lowlands. There is no liquid water on Venus, however, so its lowlands are not really seafloors, and the remaining highlands are not like the well-defined continents you see on Earth. Whereas Earth is dominated by plate tectonics, something different is happening on Venus.

The highland area Ishtar Terra, named for the Babylonian goddess of love, is about the size of Australia (■ Figure 10-4). At its eastern edge, the mountain called Maxwell Montes rises to an altitude of 12 km (7.5 mi) with the impact crater Cleopatra on its lower slopes. For comparison, Mount Everest, the tallest mountain on Earth, is 8.8 km (5.5 mi) high. Bounded by mountain ranges in the north and west, the center of Ishtar Terra is occupied by Lakshmi Planum, a great plateau about 4 km (2.5 mi) above the surrounding plains. The collapsed calderas Colette and Sacajawea suggest that Lakshmi Planum is a lava plain. The mountains bounding Ishtar Terra, including Maxwell, resemble folded mountain ranges, which suggests that limited horizontal motion in the crust as well as volcanism may have helped form the highlands.

As usual, you can learn more about other worlds by comparing them with each other and with Earth. Study **Volcanoes** on pages 194–195 and notice three important ideas plus two new terms:

1. There are two main types of volcanoes found on Earth. *Composite volcanoes* are associated mostly with plate boundaries, and *shield volcanoes* are associated with hot spots that are not related to plate boundaries.
2. Volcanoes on Venus and Mars, unlike on Earth, all have the shape of shield volcanoes, the kind produced by hot-spot volcanism rather than by plate tectonics.

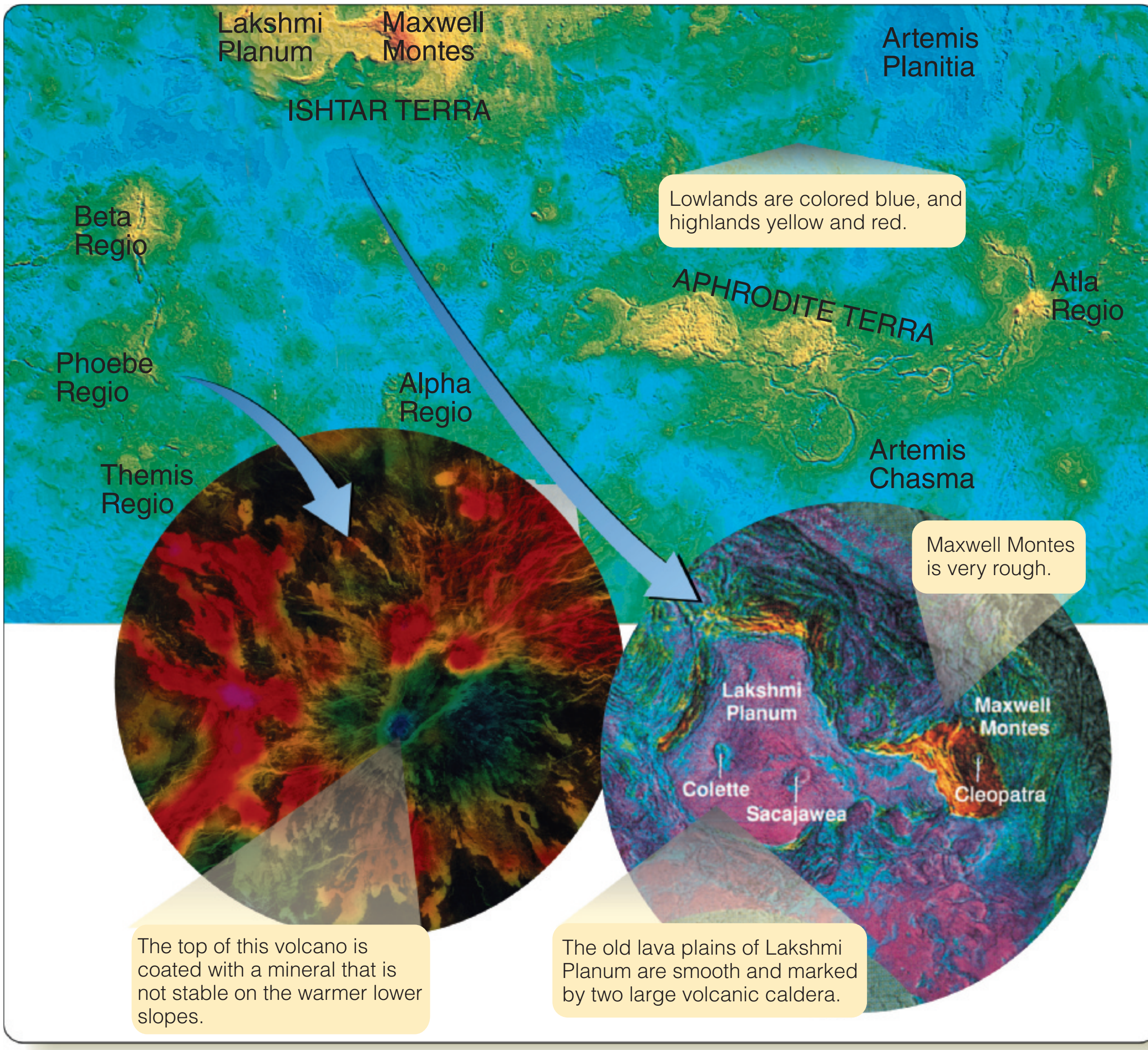


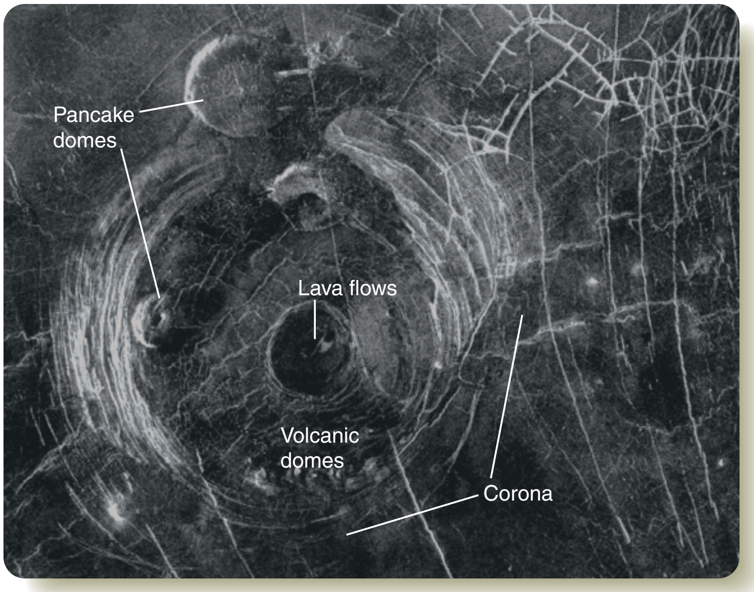
Figure 10-4

Three radar maps showing different aspects of Venus's surface. The main radar map at top shows elevation for most of the planet, omitting the polar areas. The inset map at left shows an electrical property of surface minerals related to chemical composition. The detailed map of Maxwell Montes and Lakshmi Planus inset at right is colored to show degree of terrain roughness, with purple smooth and orange rough.

3 Some volcanoes on Venus and Mars are very large. They have grown to great sizes because of repeated eruptions at the same place in the crust. This is also evidence that neither Venus nor Mars has been dominated by horizontal plate tectonics like Earth's.

Many features on Venus testify to its volcanic history. Radar maps reveal shield volcanoes and long lava channels. Other volcanic features include the **coronae,** circular bulges up to 2600 km (1600 mi) in diameter bordered by fractures, volcanoes, and lava flows (Figure 10-5). Coronae appear to be produced by rising convection currents of molten magma that push up under the crust. When the magma withdraws, the crust sinks back, but the circular fractures mark the edge of the corona.

No astronaut has ever stood on Venus, but a few spacecraft landed on the surface and survived the heat and pressure for a few hours. Some of those spacecraft analyzed nearby rocks and snapped a few photographs (Figure 10-6). The surface rocks on Venus are dark gray basalts much like those in Earth's ocean floors. This is further evidence that volcanism is important on Venus. While you are thinking about volcanoes, you can correct a **Common Misconception.** The molten rock that emerges from volcanoes comes from pockets of melted rock in the upper mantle and lower crust and not from a planet's molten core.



■ **Figure 10-5**

Radar image of a volcanic region on Venus. Aine Corona, about 200 km (125 mi) in diameter, is marked by faults, lava flows, small volcanic domes, and pancake domes of solidified lava.

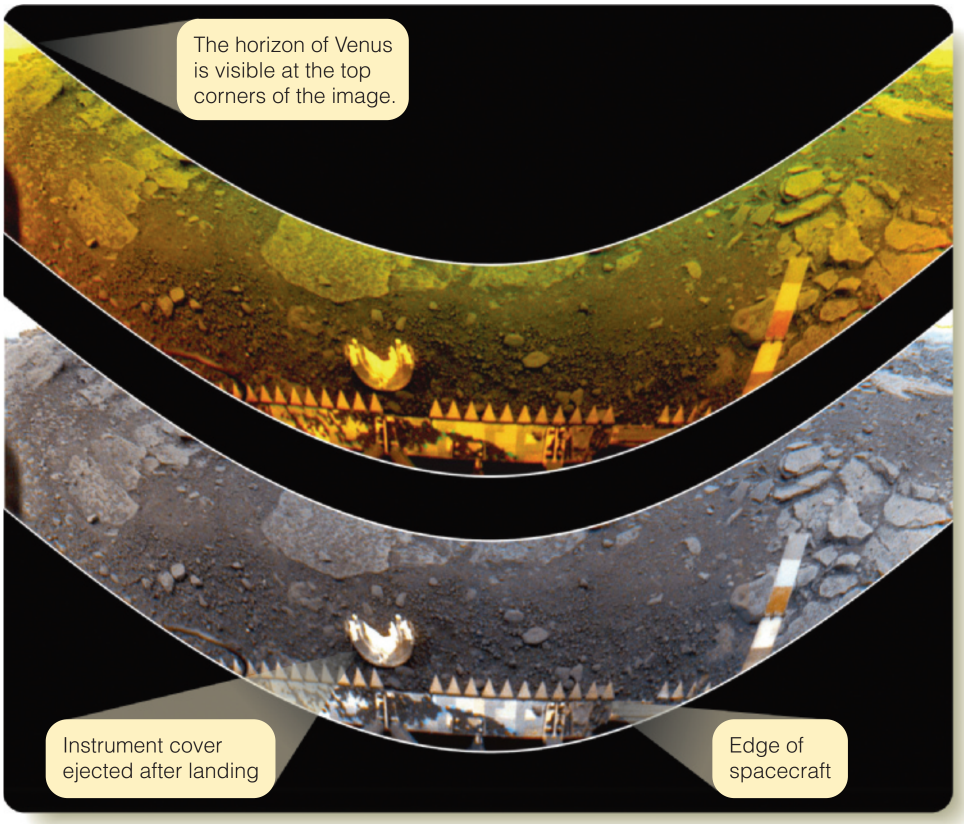
Radar images show that Venus is also marked by numerous craters (■ Figure 10-7). The atmosphere protects the surface from smaller meteorites that would produce craters smaller than 3 km (2 mi) in diameter. Larger meteorites penetrate the atmosphere and have formed about 10 percent as many craters on Venus as on the maria of Earth's moon. The number of craters shows that the crust is not as ancient as the lunar maria but also not as young as Earth's active surface. The average age of the surface of Venus is estimated from crater counts to be roughly half a billion years. Geologic processes are not renewing the surface of Venus as rapidly as Earth's surface, but no heavily cratered terrain or large impact basins remain on Venus from the heavy bombardment era.

The History of Venus

To tell the story of Venus you must draw together all the evidence and find hypotheses to explain two things, the thick carbon dioxide atmosphere and the peculiar geology.

Calculations show that Venus and Earth should have outgassed about the same amount of carbon dioxide, but Earth's oceans have dissolved most of Earth's carbon dioxide and converted it to sediments such as limestone. If all of Earth's carbon were dug up and converted back to carbon dioxide, our atmosphere would be about as dense as the air on Venus and composed mostly of carbon dioxide, like Venus's atmosphere. This suggests that the main difference between Earth and Venus is the lack of water on Venus that would have removed carbon dioxide from the atmosphere. There is evidence that Venus had oceans when it was young, but being closer to the sun it was warmer, and the carbon dioxide in the atmosphere created a greenhouse effect that made the planet even warmer. That process could have vaporized any oceans that did exist and reduced the ability of the planet to purge its atmosphere of carbon dioxide. As more carbon dioxide was outgassed, the greenhouse effect grew even more severe. Thus, Venus was trapped in what is called a runaway greenhouse effect.

The intense heat at the surface may have affected the geology of Venus by making the crust more flexible so that it was unable to break into moving plates as on Earth. There are no signs of real global plate tectonics on Venus but rather evidence that convection currents below the crust are deforming the crust to create coronae and push up mountains such as Maxwell. Other mountains, like those around Ishtar Terra, appear to be folded



■ **Figure 10-6**

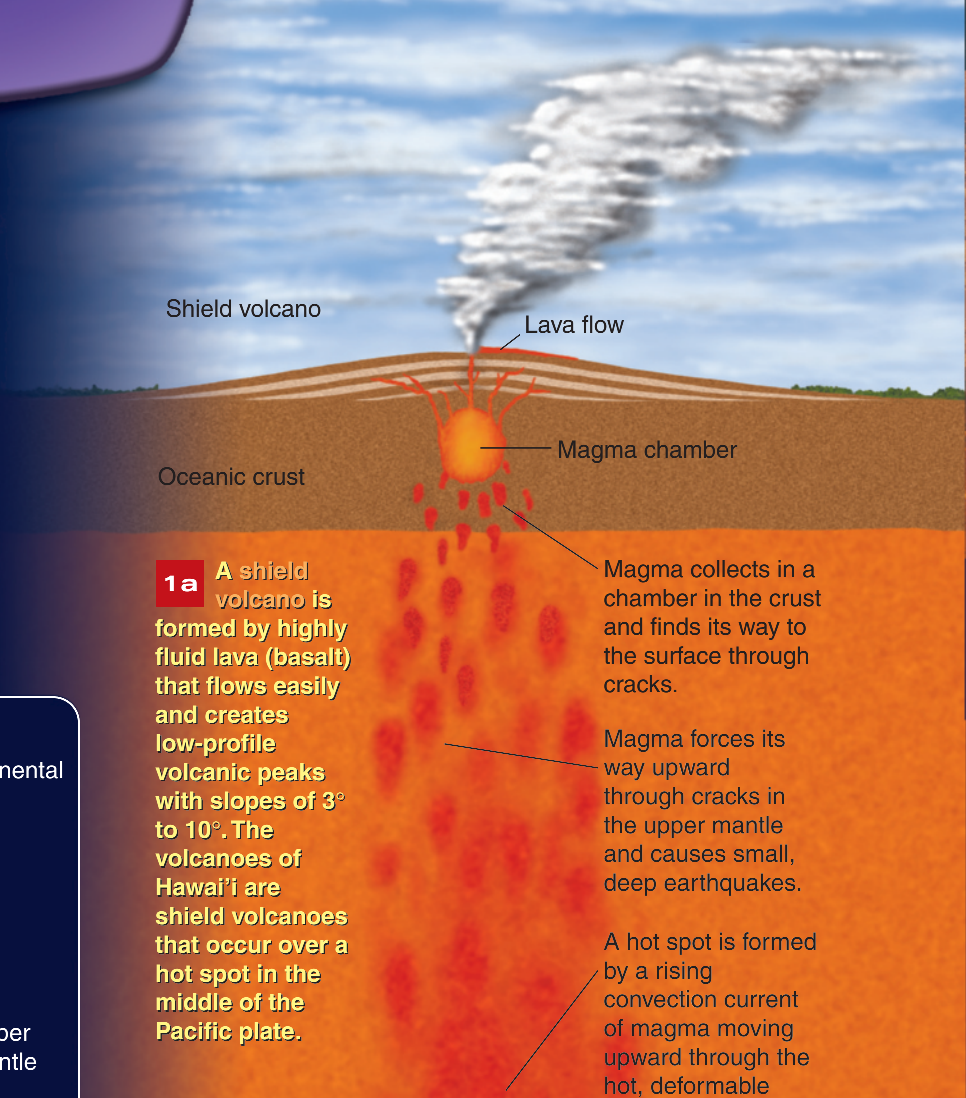
The *Venera 13* lander touched down on Venus in 1982 and carried a camera that swiveled from side to side to photograph the surface. The orange glow is produced by the thick atmosphere; when that is corrected digitally, you can see that the rocks are dark gray. Isotopic analysis suggests they are basalts.

Volcanoes

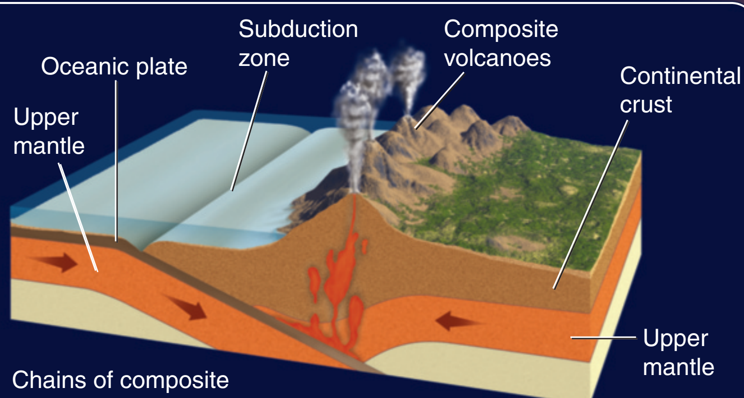
1 Molten rock (magma) is less dense than the surrounding rock and tends to rise. Where it bursts through Earth's crust, you see volcanism. The two main types of volcanoes on Earth provide good examples for comparison with those on Venus and Mars.

On Earth, **composite volcanoes** form above subduction zones where the descending crust melts and the magma rises to the surface. This forms chains of volcanoes along the subduction zone, such as the Andes along the west coast of South America.

Magma rising above subduction zones is not very fluid, and it produces explosive volcanoes with sides as steep as 30°.



1a **A shield volcano is formed by highly fluid lava (basalt) that flows easily and creates low-profile volcanic peaks with slopes of 3° to 10°. The volcanoes of Hawai'i are shield volcanoes that occur over a hot spot in the middle of the Pacific plate.**



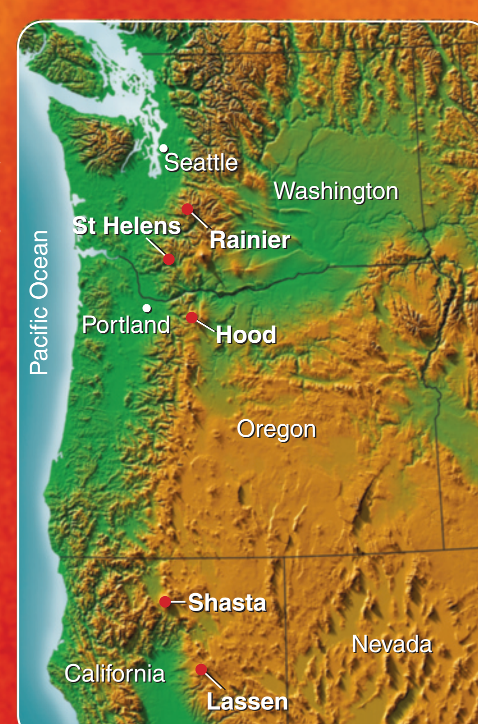
Chains of composite volcanoes are not found on Venus or Mars, which is evidence that subduction and plate motion does not occur on those worlds.

Based on *Physical Geology*, 4th edition, James S. Monroe and Reed Wicander, Wadsworth Publishing Company. Used with permission.

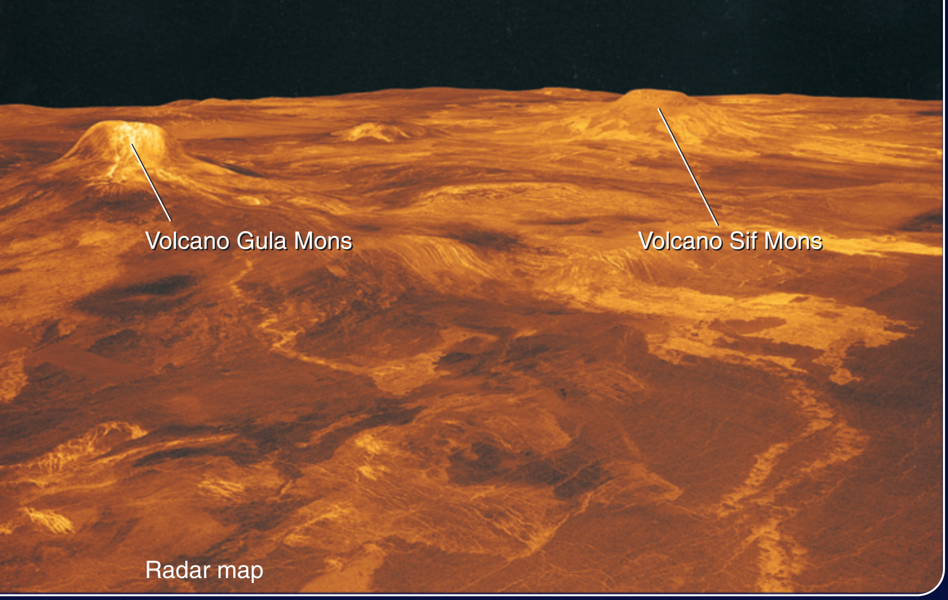
Mount St. Helens exploded northward on May 18, 1980, killing 63 people and destroying 600 km^2 (230 mi^2) of forest with a blast of winds and suspended rock fragments that moved as fast as 480 km/hr (300 mph) and had temperatures as hot as 350°C (660°F). Note the steep slope of this composite volcano.

USGS

The Cascade Range composite volcanoes are produced by an oceanic plate being subducted below North America and partially melting.



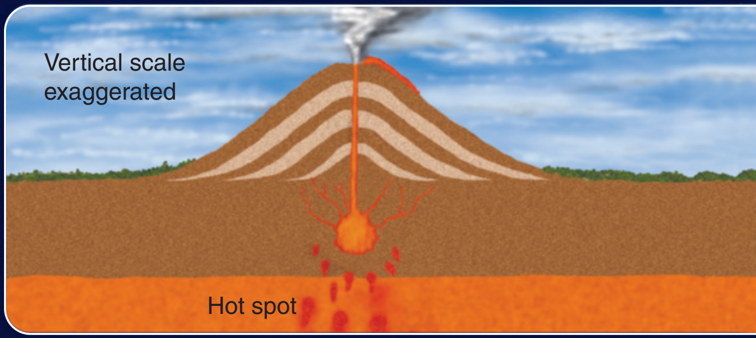
Line art on this page © Cengage Learning 2014



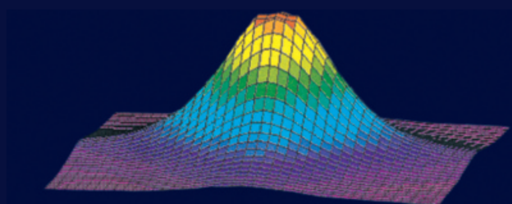
NASA

2 Volcanoes on Venus are shield volcanoes. They appear to be steep sided in some images created from Magellan radar maps, but that is because the vertical scale has been exaggerated to enhance detail. The volcanoes of Venus are actually shallow-sloped shield volcanoes.

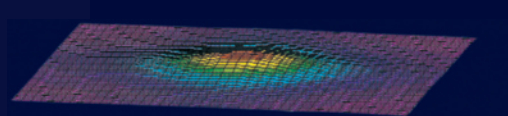
3 Volcanism over a hot spot results in repeated eruptions that build up a shield volcano of many layers. Such volcanoes can grow very large.



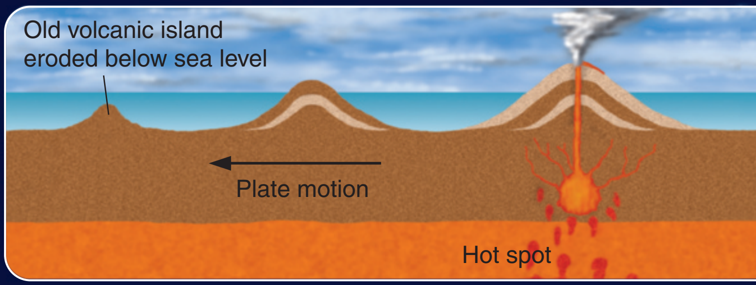
2a This computer model of a mountain with the vertical scale magnified 10 times appears to have steep slopes such as those of a composite volcano.



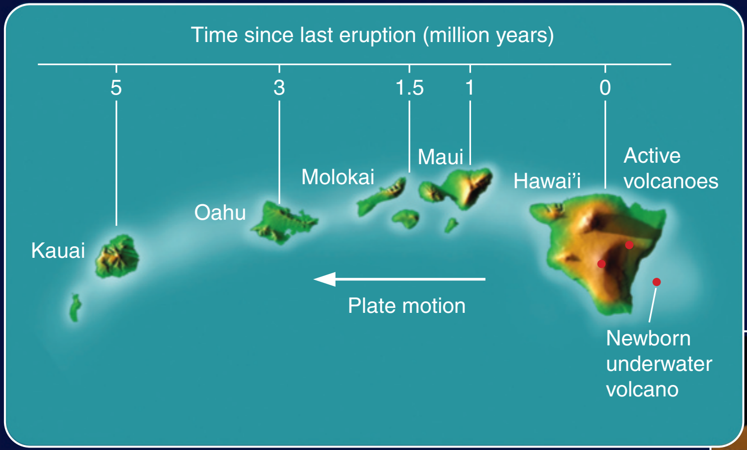
A true profile of the computer model shows the mountain has very shallow slopes typical of shield volcanoes.



Mike Seeds



If the crustal plate is moving, magma generated by the hot spot can repeatedly penetrate the crust to build a chain of volcanoes. Only the volcanoes over the hot spot are active. Older volcanoes slowly erode away. Such volcanoes cannot grow large because the moving plate carries them away from the hot spot.

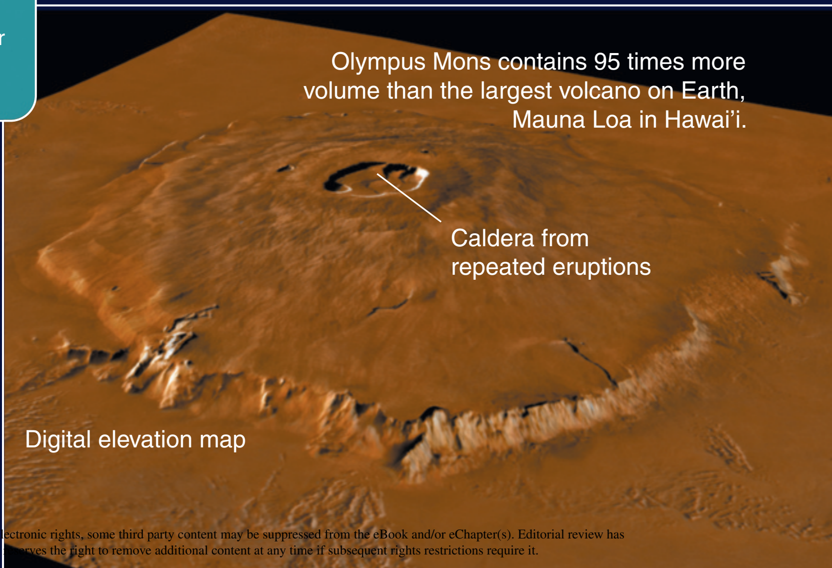


3a The volcanoes that make up the Hawaiian Islands as shown at left have been produced by a hot spot poking upward through the middle of the moving Pacific plate.

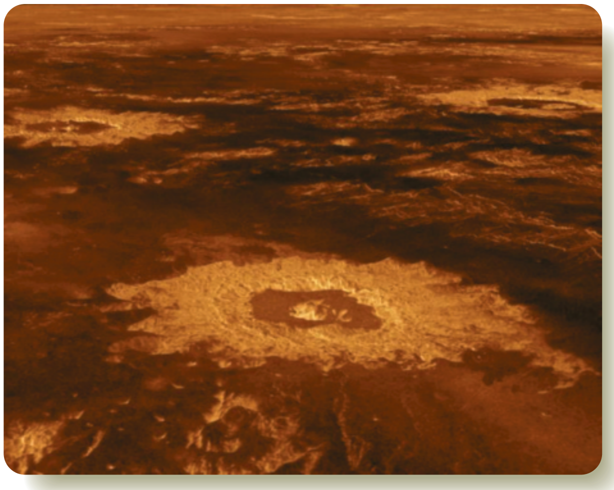
NASA

3b The plate moves about 9 cm/yr and carries older volcanic islands northwest, away from the hot spot. The volcanoes cannot grow extremely large because they are carried away from the hot spot. New islands form to the southeast over the hot spot.

Olympus Mons at right is the largest volcano on Mars. It is a shield volcano 25 km (16 mi) high and 700 km (440 mi) in diameter at its base. Its vast size is evidence that the crustal plate must have remained stationary over the hot spot. This is evidence that Mars has not had plate tectonics.



Line art on this page © Cengage Learning 2014



NASA/JPL

■ **Figure 10-7**

Radar images of impact craters on Venus: Crater Howe is 37 km (23 mi) in diameter. Craters in the background are 47 km (29 mi) and 63 km (39 mi) in diameter, respectively.

mountains caused by limited horizontal motions in the crust, driven perhaps by convection in the mantle.

The small number of craters on the surface of Venus indicates that the entire crust has been replaced within the last half-billion years or so. The resurfacing may have occurred in a planetwide overturning as the old crust broke up and sank, and lava flows created a new crust. Hypothesizing such drama may not be necessary, however. Models of the climate on Venus show that an outburst of volcanism could increase the greenhouse effect and drive the surface temperature up by as much as 100°C. This could further soften the crust, increase the volcanism, and push the planet into a resurfacing episode. This type of catastrophe may happen periodically on Venus, or the planet may have had a single, geologically recent resurfacing event. In either case, un-Earthly Venus may eventually reveal more about how our own world works.

SCIENTIFIC ARGUMENT

What evidence indicates that Venus does not have plate tectonics?
This argument must cite evidence and use comparison. On Earth, plate tectonics is identifiable by the worldwide network of faults, subduction zones, volcanism, and folded mountain chains that outline the plates. Although some of these features are visible on Venus, they do not occur in a planetwide network that outlines multiple plates. Volcanism is widespread, but folded mountain ranges occur in only a few places. Rather than being dominated by the horizontal motion of rigid crustal plates, Venus may have a more flexible crust dominated instead by vertical tectonics, for example, rising plumes of molten rock that strain the crust to produce coronae or break through to form volcanoes and lava flows.

Earth and Venus are sibling worlds in some ways, but in other ways they seem to be no more than distant cousins. Now build an argument to compare atmospheres. **Why isn't Earth's atmosphere like that of Venus?**

10-3 Mars

MARS IS A MEDIUM-SIZED WORLD about half the diameter of Earth (**Celestial Profile 6**, p. 197). The surface is old, cratered, and marked by volcanoes, but as you explore, watch for evidence that water once flowed there.

The Atmosphere of Mars

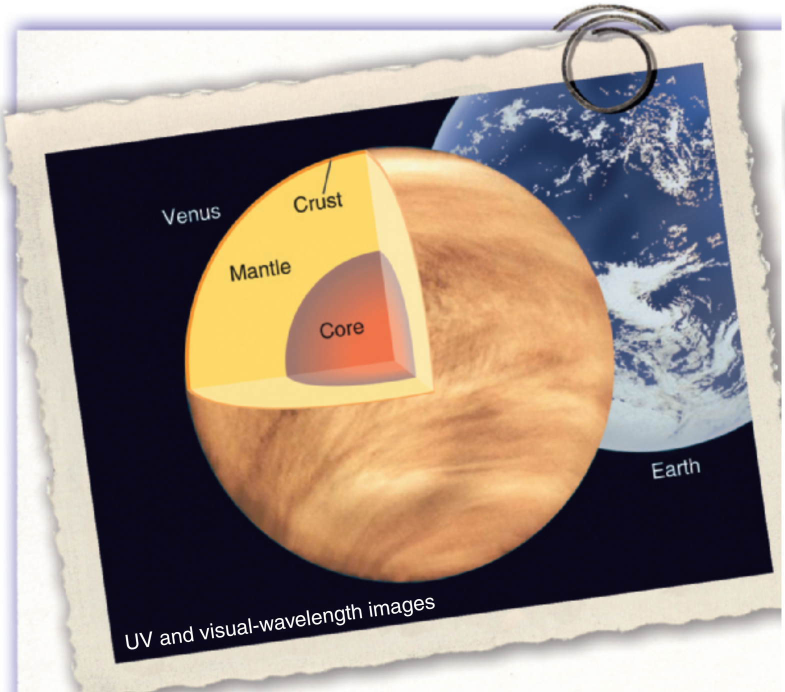
The Martian air contains 95 percent carbon dioxide, 3 percent nitrogen, and 2 percent argon. That is much like the chemical composition of the air on Venus, but the Martian atmosphere is very thin, less than 1 percent as dense as Earth's atmosphere, one ten-thousandth as dense as Venus's atmosphere.

There is very little water in the Martian atmosphere, and the polar caps are composed of frozen water ice coated over by frozen carbon dioxide ("dry ice"). As summer comes to a Martian hemisphere, planetary scientists observe the carbon dioxide in that polar cap turning from solid to vapor and adding carbon dioxide to the atmosphere, while winter in the opposite hemisphere is freezing carbon dioxide out of the atmosphere and adding it to that polar cap.

Liquid water cannot survive on the surface of Mars because the air pressure is too low. Any liquid water would immediately boil away; and if you stepped out of a spaceship on Mars without your spacesuit, your body heat would make your blood boil. Whatever water is present on Mars must be frozen in the polar caps or in the form of **permafrost** within the soil.

Although the present atmosphere of Mars is very thin, you will see evidence that the climate once permitted liquid water to flow over the surface, so Mars must have once had a thicker atmosphere. As a Terrestrial planet, it should have outgassed significant amounts of carbon dioxide, nitrogen, and water vapor, but because it was small it could not hold on to its gases. The escape velocity on Mars is only 5 km/s, less than half of Earth's, so it was easier for rapidly moving gas molecules to escape into space. Another factor is the temperature of a planet. If Mars had been colder, the gas molecules in its atmosphere would have been traveling more slowly and would not have escaped as easily. You can see this in ■ Figure 10-8 that plots the escape velocity of each planet versus the temperature of the region from which molecules would escape. For Earth, the temperature is that of the upper atmosphere. For Mercury, the temperature is that of the hot rocky surface. Clearly, small worlds cannot keep atmospheric gases easily.

A further problem is that Mars has no ozone layer to protect its atmosphere from ultraviolet radiation. (Sunbathing on Mars would be a fatal mistake.) The ultraviolet photons can break atmospheric molecules up into smaller fragments, which escape more easily. For example, water can be broken up into hydrogen and oxygen. Thus, Mars is large enough to have had a substantial atmosphere when it was young, and may have had water falling as rain and collecting in rivers and lakes. It gradually lost much of its



© Cengage Learning 2014; image: NASA

Venus's diameter is 95 percent that of Earth. Its atmosphere is perpetually cloudy, and its surface is hot enough to melt lead. Whether it has a liquid metal core like Earth's is currently a matter of conjecture.

Celestial Profile 5: Venus

Motion:

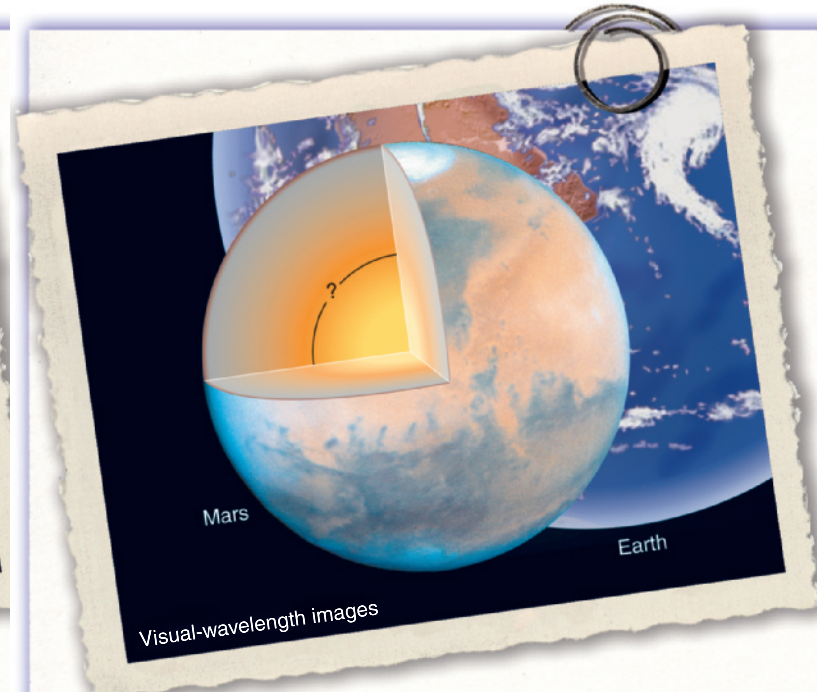
Average distance from the sun	0.723 AU (1.08×10^8 km)
Eccentricity of orbit	0.007
Inclination of orbit to ecliptic	3.4°
Orbital period	0.6152 y (224.70 d)
Period of rotation	243.0 d
Inclination of equator to orbit	177.3° (retrograde rotation)

Characteristics:

Equatorial diameter	1.21×10^4 km (0.949 $D_\oplus$)
Mass	4.87×10^{24} kg (0.815 $M_\oplus$)
Average density	5.20 g/cm^3 (4.2 g/cm^3 uncompressed)
Surface gravity	0.90 Earth gravity
Escape velocity	10.5 km/s (0.92 $V_\oplus$)
Surface temperature	470°C (880°F)
Albedo (cloud tops)	0.90
Oblateness	0

Personality Point:

Venus is named for the Roman goddess of love, perhaps because the planet often shines so beautifully in the evening or dawn sky. In contrast, the ancient Maya identified Venus as their war god Kukulkan and sacrificed human victims to the planet when it rose in the dawn sky.



© Cengage Learning 2014; image: NASA

Mars has half the diameter of Earth and probably retains some internal heat, but the size and composition of its core are not well known.

Celestial Profile 6: Mars

Motion:

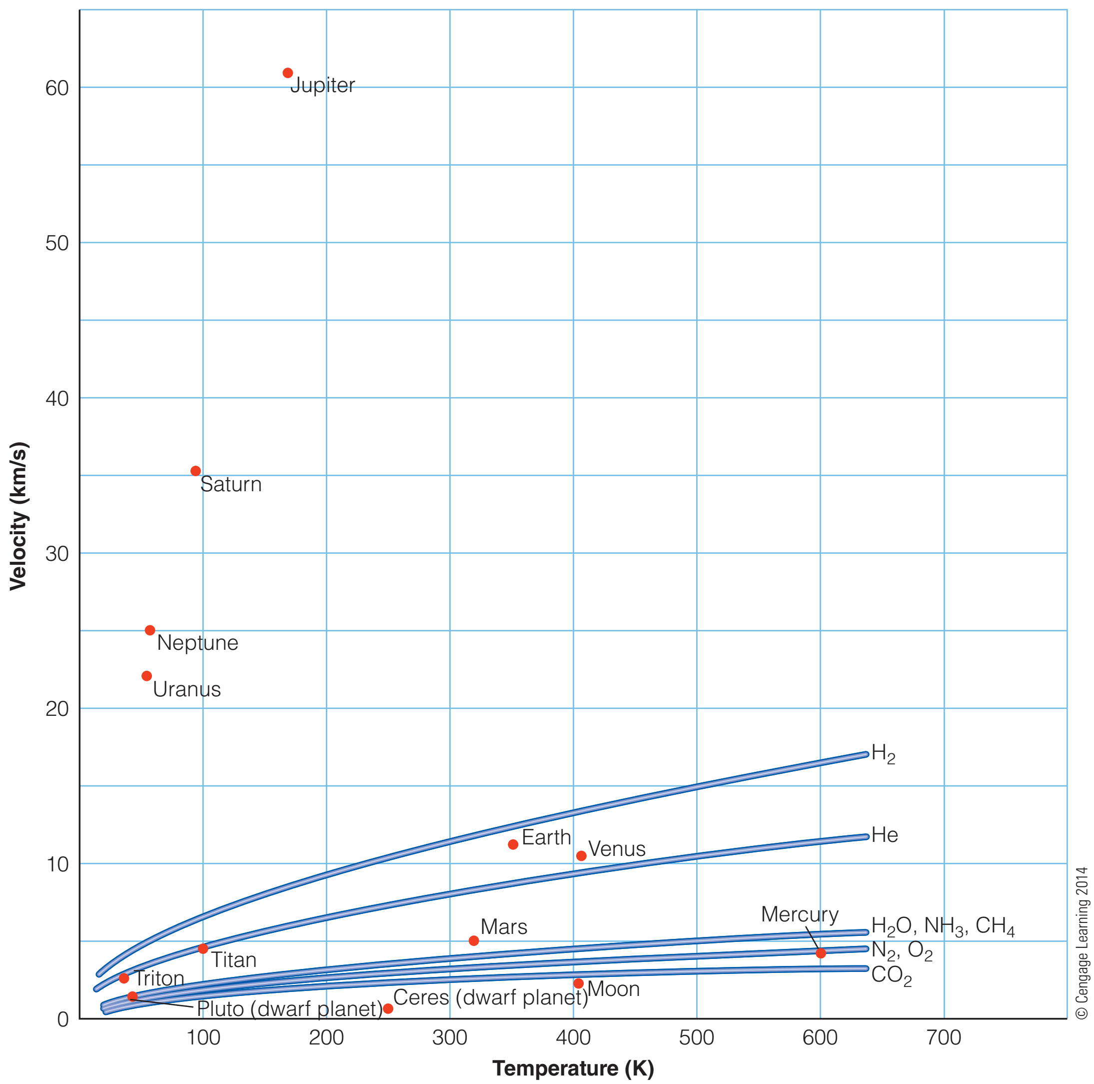
Average distance from the sun	1.52 AU (2.28×10^8 km)
Eccentricity of orbit	0.093
Inclination of orbit to ecliptic	1.8°
Orbital period	1.881 y (687.0 d)
Period of rotation	24.62 h
Inclination of equator to orbit	25.2°

Characteristics:

Equatorial diameter	6.79×10^3 km (0.531 $D_\oplus$)
Mass	6.42×10^{25} kg (0.107 $M_\oplus$)
Average density	3.93 g/cm^3 (3.3 g/cm^3 uncompressed)
Surface gravity	0.38 Earth gravity
Escape velocity	5.0 km/s (0.45 $V_\oplus$)
Surface temperature	$-140°$ to 15°C ($-220°$ to 60°F)
Average albedo	0.25
Oblateness	0.009

Personality Point:

Mars is named for the god of war. Minerva was the goddess of defensive war, but Bullfinch's *Mythology* refers to Mars's "savage love of violence and bloodshed." You can see the planet glowing reddish orange from Earth, reminiscent of blood to cultures throughout history, because of iron oxides in its soil.



■ Figure 10-8

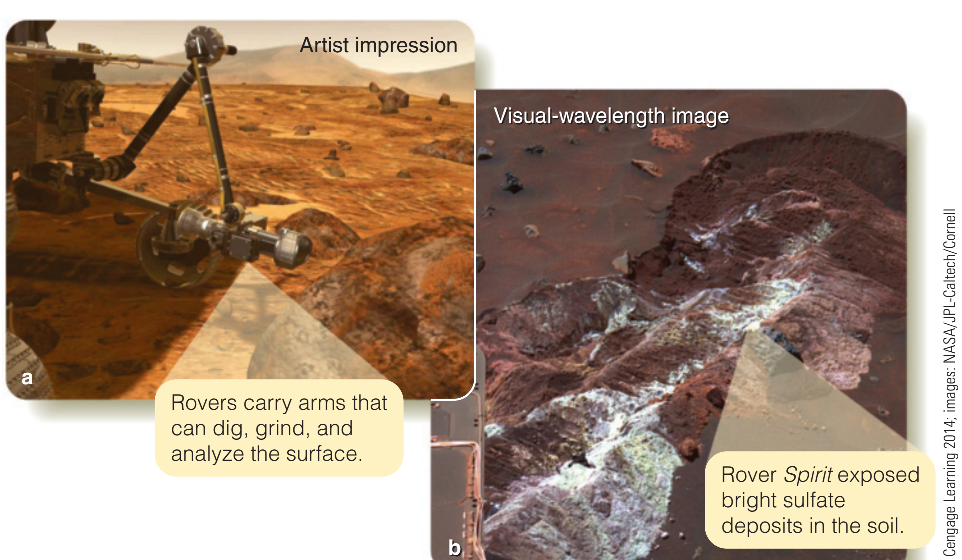
This plot shows the ability of planets to retain atmospheres. Dots represent the escape velocity and temperature of various solar system bodies. The lines represent the typical highest velocities of gas molecules of various masses. The Jovian planets have high escape velocities and can hold on to even the lowest-mass molecules. Mars can hold only the more massive molecules, and the moon has such a low escape velocity that all gas molecules can escape.

atmosphere and is now a cold, dry world.

Exploring the Surface of Mars

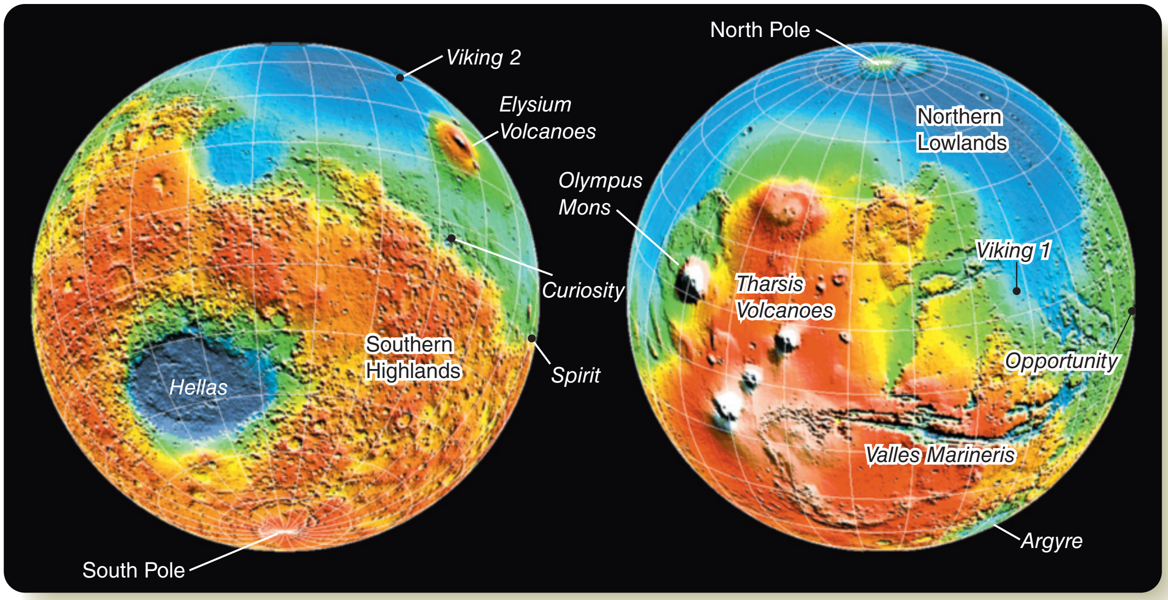
If you ever visit another world, Mars may be your best choice. You will need a heated, pressurized spacesuit, but Mars is much more hospitable than the moon or Venus. It is also more interesting, with weather, complex geology, and signs that water once flowed over its surface. You might even hope to find traces of ancient life hidden in the rocks.

Spacecraft have been visiting Mars for almost 40 years. Two *Viking* landers first touched down in 1976. Rovers *Spirit* and *Opportunity* landed in 2004 carrying instruments to explore the rocky surface (■ Figure 10-9); *Opportunity* was still operating and in communication with Earth eight years later. The *Phoenix* robot laboratory landed in the north polar region in 2008. The sophisticated rover laboratory *Curiosity* landed and began its exploration of Mars in summer 2012.



■ Figure 10-9

(a) Rovers *Spirit* and *Opportunity* were directed from Earth to move across the surface of Mars, explore features, dig in the soil, grind the surfaces of rocks to expose their interiors, and make spectroscopic analyses. (b) Rover *Spirit's* discovery of sulfate deposits in the soil confirms other evidence that a body of salty water once covered the area and evaporated, leaving the sulfates behind.



© Cengage Learning 2014; image: NASA

Figure 10-10

These hemisphere maps of Mars are color coded to show elevation. The northern lowlands lie about 4 km (2.5 mi) below the southern highlands. Volcanoes are very high (white), and the giant impact basins, Hellas and Argyre, are low. Note the depth of the canyon Valles Marineris. The two *Viking* spacecraft landed on Mars in 1976, *Pathfinder* in 1997, and *Spirit* plus *Opportunity* in 2004.

Images made by a series of orbiting probes show that the southern hemisphere of Mars is a heavily cratered highland region estimated to be at least 2 to 3 billion years old. The northern hemisphere is mostly a much younger lowland plane with few craters (■ Figure 10-10). This lowland plane may have been smoothed by lava flows, but growing evidence suggests that it was once filled with an ocean, a controversial hypothesis discussed in the next section.

Volcanism on Mars is dramatically evident in the Tharsis region, a highland region of volcanoes and lava flows bulging 10 km (6 mi) above the surrounding surface. A similar uplifted volcanic plain, the Elysium region, is more heavily cratered and eroded and appears to be older than the Tharsis bulge. The lack of many impact craters suggests that some volcanoes have been active within the last few hundred million years.

All of the volcanoes on Mars are shield volcanoes, which you learned earlier are produced by hot spots penetrating upward through the crust and are not related to plate tectonics. Olympus Mons, the largest volcano on Mars, is 700 km (440 mi) in diameter at its base and rises 25 km (16 mi) high. The largest volcano on Earth, Mauna Loa in Hawaii, rises only 10 km (6 mi) above its base on the seafloor. Mauna Loa is so heavy that it has sunk into Earth's crust, producing an undersea moat around its base. In contrast, Olympus Mons, more than twice as high, has no moat and is supported entirely by the Martian crust (■ Figure 10-11). Evidently, the crust of Mars is stronger and thicker than Earth's.

The size of Olympus Mons also provides clear evidence that plate tectonics has not been significant on Mars. On Earth, shield volcanoes such as those that formed the Hawaiian Islands occur over rising currents of hot material in the mantle. Because the plate moves, the hot material heats the crust in a string of locations and forms a chain of volcanoes instead of a single large feature. The Hawaiian Islands are merely the most recent of a series of volcanic islands called the Hawaiian–Emperor Island chain (see page 172), which stretches nearly 3800 km (2400 mi) across the Pacific Ocean floor. In contrast, a lack of plate motion on Mars has allowed rising currents of magma to heat the crust repeatedly at the locations of Olympus Mons and other volcanoes, building huge volcanic cones much larger than any on Earth.

When the crust of a planet is strained, it may break, producing faults and rift valleys. Near the Tharsis region is a great valley, Valles Marineris (Figure 10-11b), named after the *Mariner 9* spacecraft that first photographed it. This valley is a block of crust that has dropped downward along parallel faults. Erosion and landslides have further modified the valley into a great canyon stretching almost one-fifth of the way around the planet. It is four times deeper, nearly ten times wider, and over ten times longer than the Grand Canyon. The number of craters in the valley indicates that it is 1 to 2 billion years old, placing its origin sometime before the end of the most active volcanism in the Tharsis region.

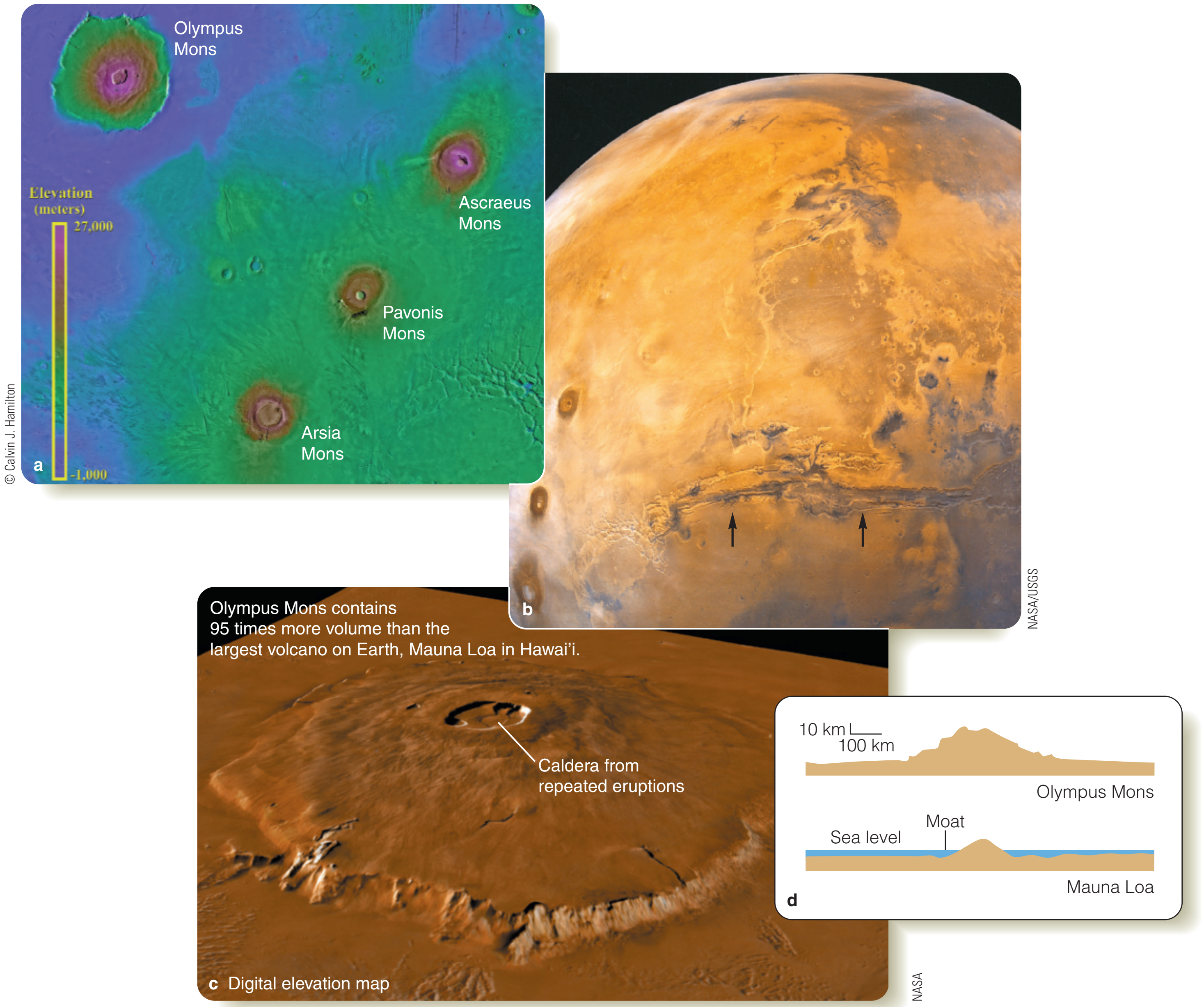


Figure 10-11

(a) High volcanoes and deep canyons mark the surface of Mars. Olympus Mons, a shield volcano, is much larger than the largest volcano on Earth. In this false-color image, three other volcanoes are visible in the Tharsis region. (b) The three Tharsis volcanoes are also visible in this *Viking* spacecraft photo along with the canyon Valles Marineris (indicated by arrows), which stretches as far as the distance from New York to Los Angeles. (c) Map of Olympus Mons, the largest volcano on Mars. The vast size of this shield volcano is evidence that the crust of Mars must have remained stationary over the hot spot, and Mars has not had horizontal plate tectonics. (d) Comparison of the profiles of shield volcano Olympus Mons on Mars with shield volcano Mauna Loa on Earth. Mauna Loa is the largest single mountain mass on our planet.

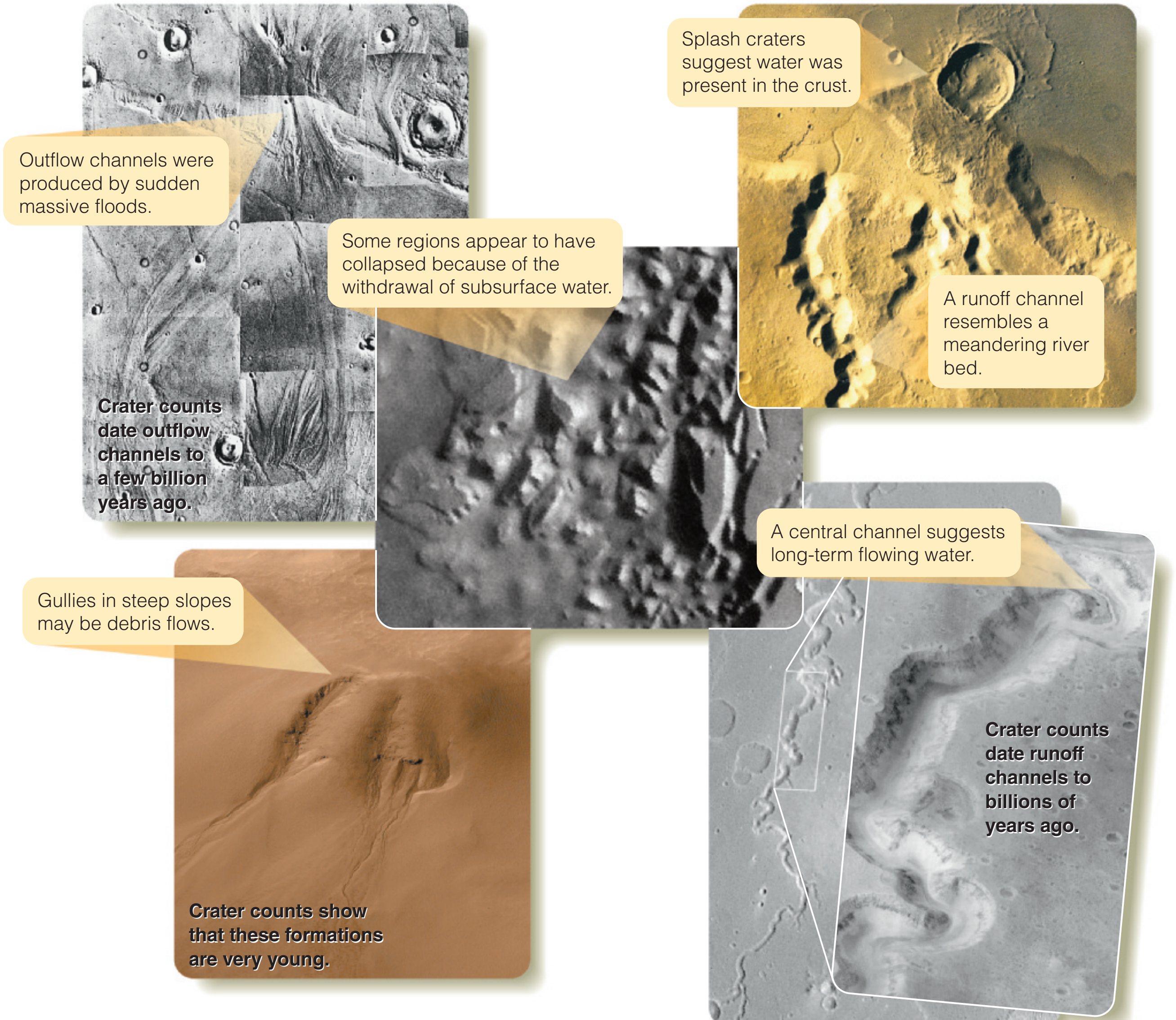


Figure 10-12

These visual-wavelength images made by the *Viking* orbiters and Mars Global Surveyor show some of the features that suggest liquid water on Mars. Outflow channels and runoff channels are old, but some gullies may be quite recent.

Before you can tell the story of Mars, you must consider a difficult issue—water. How much water has Mars had, how much has been lost, and how much remains?

Searching for Water on Mars

You would hardly expect water on the surface of Mars. It is a cold, dry desert world. However, observations from orbiting spacecraft have revealed landforms that suggest there was once water on Mars, and rovers on the surface have turned up further evidence of past water.

In 1976, the two *Viking* spacecraft reached orbit around Mars and photographed its surface. Those photos revealed two kinds of water-related features. **Outflow channels** appear to have been cut by massive floods carrying as much as 10,000 times the volume of water flowing down the Mississippi River. In a matter of hours or days, such floods swept away geological features and left scarred land such as that shown in ■ Figure 10-12. In contrast, **valley networks** look like meandering riverbeds with sandbars, deltas, and tributaries typical of streams that flowed for extended periods of time. The number of craters on top of these features reveals that they are typically over 3 billion years old.

Images made from orbit also show regions of jumbled terrain, suggesting that subsurface ice may have melted and drained away. Gullies leading down slope suggest water seeping from underground sources. The terrain at the edges of the northern lowlands resembles Earth shorelines, and some scientists suspect that the northern lowlands were filled with an ocean roughly

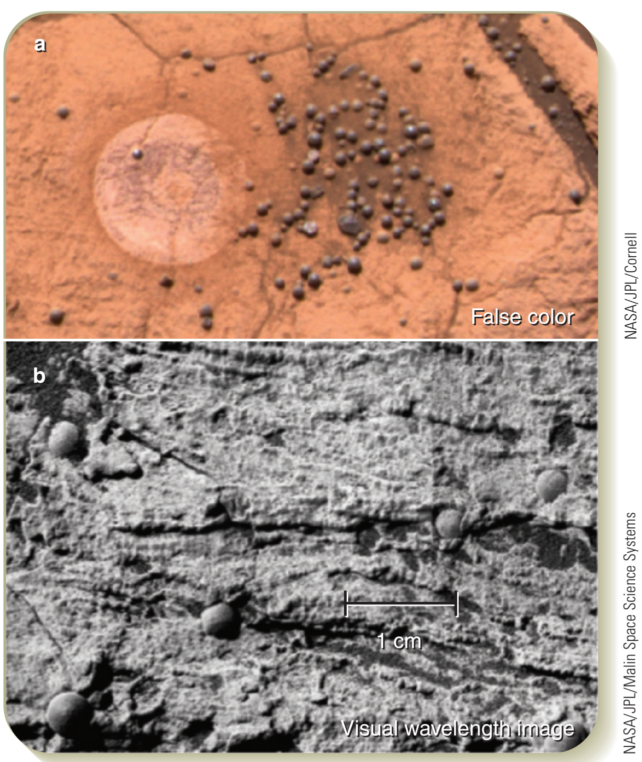
3 billion years ago. Look again at Figure 10-10, where the lowlands have been color-coded blue, and notice the major outflow channels leading from the highlands into the lowlands northwest of the *Viking 1* landing site and southeast of the *Pathfinder* landing site, like rivers flowing into an ocean. Careful measurements of the location of the hypothetical ocean's shoreline indicate that it might originally have been an enormous impact basin.

Spacecraft in orbit around Mars have used remote instruments to detect large amounts of water frozen in the soil. A radar study has found frozen water extending at least a kilometer beneath both polar caps.

Rovers *Spirit* and *Opportunity* were both targeted to land in areas suspected of having had water on their surfaces, and they each made important discoveries. Using close-up cameras, they found small spherical concretions of the mineral hematite (nicknamed "blueberries") that must have formed in water. In other places, they found layers of sediments with ripple marks and crossed layers showing they were deposited in moving water (■ Figure 10-13). Chemical analysis revealed minerals in the soil such as sulfates that would have been left behind when standing water evaporated.

Orbiting spacecraft have photographed layered terrain near the polar caps (■ Figure 10-14). Year after year dust accumulates on the polar caps and is then left behind in a layer when the polar caps vaporize in the spring. Over periods of thousands of years, deep layers can develop. What is significant is that orbiters have photographed newer layers oriented differently from older underlying layers, showing that the climate and wind patterns on Mars have changed repeatedly. These layers suggest that the climate on Mars may vary because of cyclic changes in the rotation and orbital revolution of the planet. Recall from Chapter 2 that Earth is affected by such cycles.

Mars has water, but it is hidden. The climate has changed time after time, but the atmosphere has gradually grown thinner. The oceans and lakes are gone. The last of the water on Mars is in the polar caps or frozen in the crust. Water is the first necessity of life, so the evidence for running water long ago on Mars is exciting. Someday an astronaut may scramble down an ancient Martian streambed, turn over a rock, and find a fossil.



■ Figure 10-13

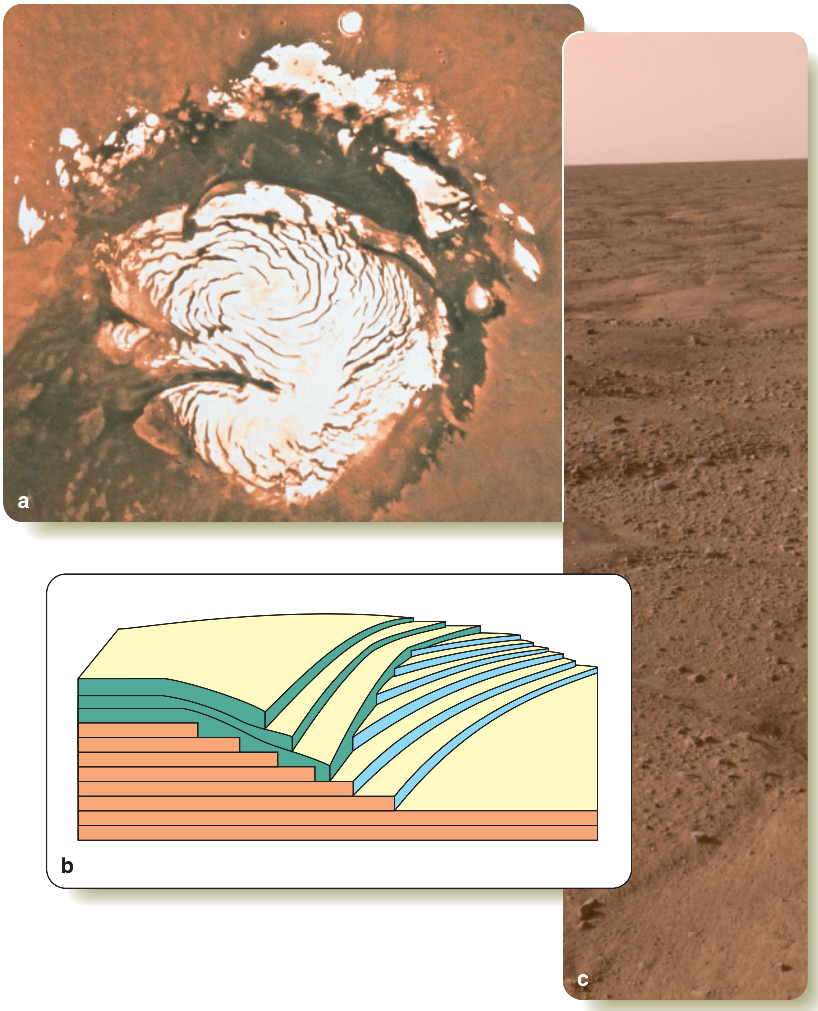
(a) Rover *Opportunity* photographed hematite concretions ("blueberries") weathered from rock. The round mark is a spot cleaned by the rover. The spheres appear to have grown as minerals collected around small crystals in the presence of water. Similar concretions are found on Earth. (b) The layers in this rock were deposited as sand and silt in rapidly flowing water. From the way the layers curve and cross each other, geologists can estimate that the water was at least ten centimeters deep. A few "blueberries" are also visible in this image.

The History of Mars

Did Mars ever have plate tectonics? Where did the water go? These fundamental questions challenge you to assemble the evidence and hypotheses for Mars and tell the story of its evolution (**How Do We Know? 10-2**).

The four-stage history of Mars is a case of arrested development. The planet began by differentiating into a crust, mantle, and core. Studies of its rotation reveal that it has a dense core. The Mars Global Surveyor spacecraft detected no planetwide magnetic field, but it did find regions of the crust with fields a bit over 1 percent as strong as Earth's. Apparently, the young Mars had a molten iron core and generated a magnetic field, which became frozen into parts of the crust. The core must have cooled quickly and shut off the dynamo effect that was producing the planet-wide field. The magnetic regions of the crust remain behind like fossils.

The crust of Mars is now quite thick, as shown by the mass of Olympus Mons, but it was probably thinner in the past. Cratering may have broken or at least weakened the crust, triggering lava flows that flooded some basins. Most of the northern hemisphere may have been a huge impact basin that was later filled with water, creating a Martian ocean that has since vanished. Mantle convection may have pushed up the Tharsis and Elysium volcanic regions and broken the crust to form Valles Marineris, but moving crustal plates never dominated Mars. There are no folded mountain ranges on Mars and no signs of plate boundaries. The large size of volcanoes on the planet is evidence that the crust does not move. Planetary scientists therefore conclude that as Mars's interior cooled its crust grew thick and immobile.



■ Figure 10-14

(a) The north polar cap of Mars is made of many regions of ice separated by narrow valleys free of ice. (b) In some regions of the polar cap, layers of ice plus dust with different orientations are superimposed, suggesting periodic changes in the Martian climate. (c) This view from the *Phoenix* lander shows the landscape of Mars's north polar plains, including polygonal cracks believed to result from seasonal expansion and contraction of ice under the surface.

At some point in the history of Mars, water was abundant enough to flow over the surface in great floods and may have filled oceans, but the age of liquid water must have ended over 3 billion years ago. The climate on Mars has changed as atmospheric gases and water were lost to space and as water was frozen into the soil as permafrost.

For Mars, the fourth stage of planetary development has been one of moderate activity and slow decline. Volcanoes may still occasionally erupt, but this medium-sized planet has lost much of its internal heat, and most volcanism occurred long ago. The atmosphere is thin, and the surface is a forbiddingly dry, cold desert.

Comparative Planetology, Once More

Venus and Mars share at least one characteristic: They have evolved since they formed and are now quite different than they were when the solar system was young. Moreover, as you have learned, planetary scientists have evidence that surface conditions on Venus, Earth, and Mars were much more similar long ago than they are now. Study **When Good Planets Go Bad** on pages 204–205 and notice four important points:

1. The difference between Venus and Earth is not in the amount of CO_2 they have outgassed but in the amount they have removed from their atmospheres. Being warmer and consequently having less liquid water on its surface sealed Venus's fate as a runaway greenhouse.
2. Venus is highly volcanic with a crust made up of lava flows that have covered over any older crust. The entire planet has been resurfaced within the last half a billion years.
3. Mars is significantly smaller than Earth, with gravity too weak to prevent much of its atmosphere from leaking away. Evidence shows that Mars once had liquid water on its surface, but much of its water has been lost to space. The low atmospheric pressure means that the only remaining water is frozen in the soil.
4. Because Mars is small, its interior cooled relatively quickly and volcanism has died down. The lack of volcanism meant the escaping atmosphere is not replenished. The planet's crust was thinner when the planet was young but has grown thick and never broke into moving plates as did Earth's crust.

The Moons of Mars

Unlike Mercury or Venus, Mars has moons. Small and irregular in shape, Phobos (28 × 23 × 20 km in diameter) and Deimos (16 × 12 × 10 km) are probably captured asteroids.

Photographs reveal a unique set of narrow, parallel grooves on Phobos (■ Figure 10-15a). Averaging 150 m (500 ft) wide and 25 m (80 ft) deep, the grooves run from Stickney, the largest crater, to an oddly featureless region on the opposite side of the satellite. One theory suggests that the grooves are deep fractures caused by the impact that formed Stickney.

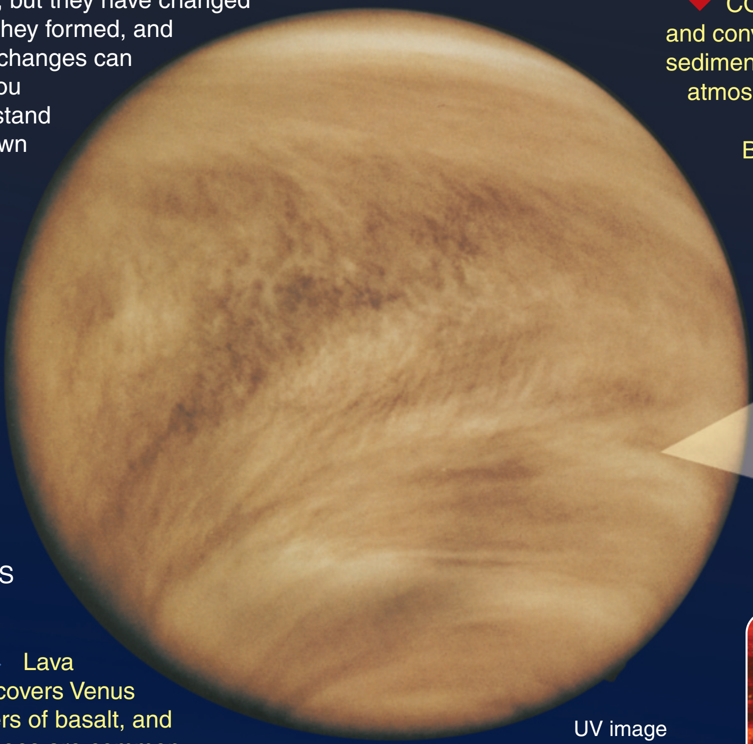
Deimos not only has no grooves but it also looks smoother because of a thicker layer of dust on its surface. This material partially fills craters and covers minor surface irregularities

Venus and Mars haven't really gone wrong, but they have changed since they formed, and those changes can help you understand your own world.

Venus: Runaway Greenhouse

1 Venus and Earth have outgassed about the same amount of CO_2, but Earth's oceans have dissolved most of Earth's CO_2 and converted it to sediments such as limestone. If all of Earth's sedimentary carbon were dug up and converted back to CO_2, our atmosphere would be much like Venus's.

Because Venus was warmer when it formed, it had little if any liquid water to dissolve CO_2, and that produced a greenhouse effect that made the planet even warmer. The planet could not purge its atmospere of CO_2, and as more was outgassed, Venus was trapped in a runaway greenhouse effect.



VENUS

UV image

NASA

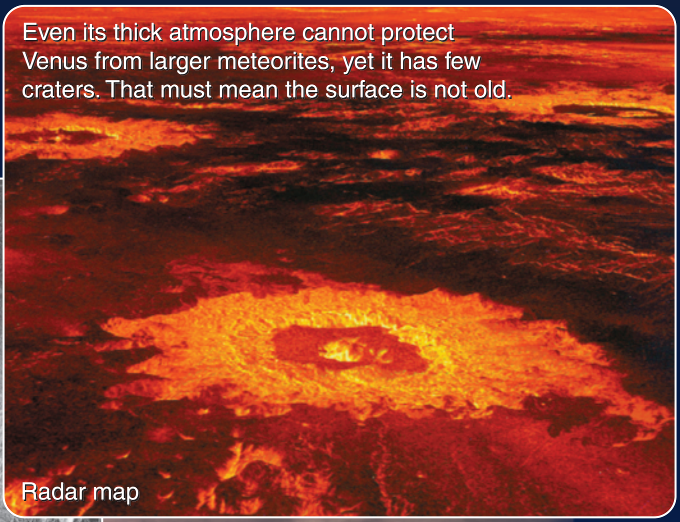
Only 0.7 AU from the sun, Venus receives almost twice the solar energy per square meter that Earth does. Moved to Venus' orbit, Earth's surface would be 50° C (90° F) hotter.

2 Lava covers Venus in layers of basalt, and volcanoes are common. Some of those volcanoes may be active right now while others are dormant. Traces of past lava flows show up in radar images including long narrow valleys cut by moving lava.

Baltis Vallis (arrows), at least 6800 km (4200 mi) long, is the longest lava flow channel in the solar system.

Radar map

NASA/JPL



Even its thick atmosphere cannot protect Venus from larger meteorites, yet it has few craters. That must mean the surface is not old.

Radar map

NASA/JPL

2a The crust of Venus must be no older than 0.5 billion years. One hypothesis is that an earlier crust broke up and sank in a sea of magma as fresh lava flows formed a new crust. Such resurfacing events might occur periodically on a hot, volcanic planet like Venus.

What could cause a resurfacing event? Models of the climate on Venus show that an outburst of volcanism could increase the greenhouse effect and drive the surface temperature up by as much as 100°C. This could soften the crust, increase the volcanism, and push the planet into a resurfacing episode. This type of catastrophe may happen periodically on Venus, or the planet may have had just one resurfacing event about half a billion years ago.

Even if bodies of water existed when Venus was young, they could not have survived long.

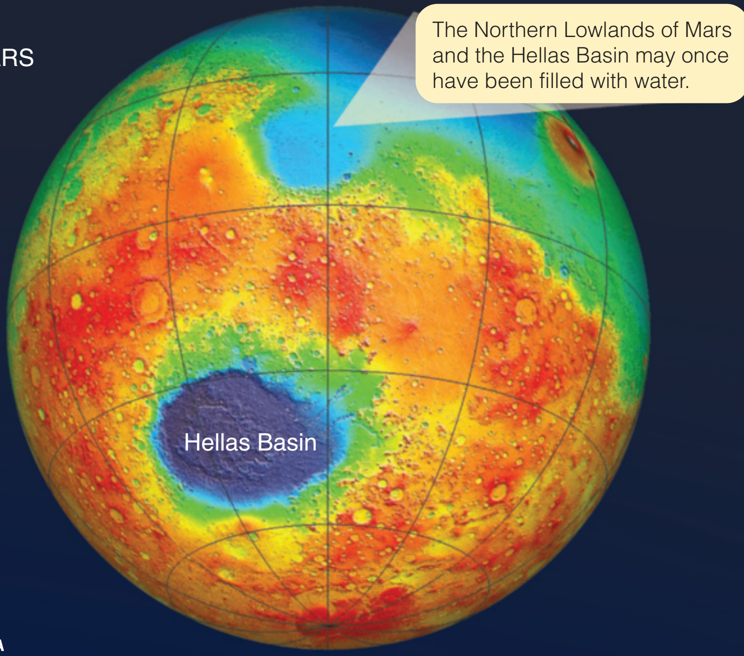
Line art on this page © Cengage Learning 2014

NASA/JPL

Mars: Runaway Refrigerator

3 When Mars was young, water was abundant enough to flow over the surface in streams and floods, and may have filled oceans. That age of liquid water ended over three billion years ago. The climage on Mars has changed as atmospheric gases and water were lost to space and as water was frozen into the soil as permafrost.

MARS



NASA

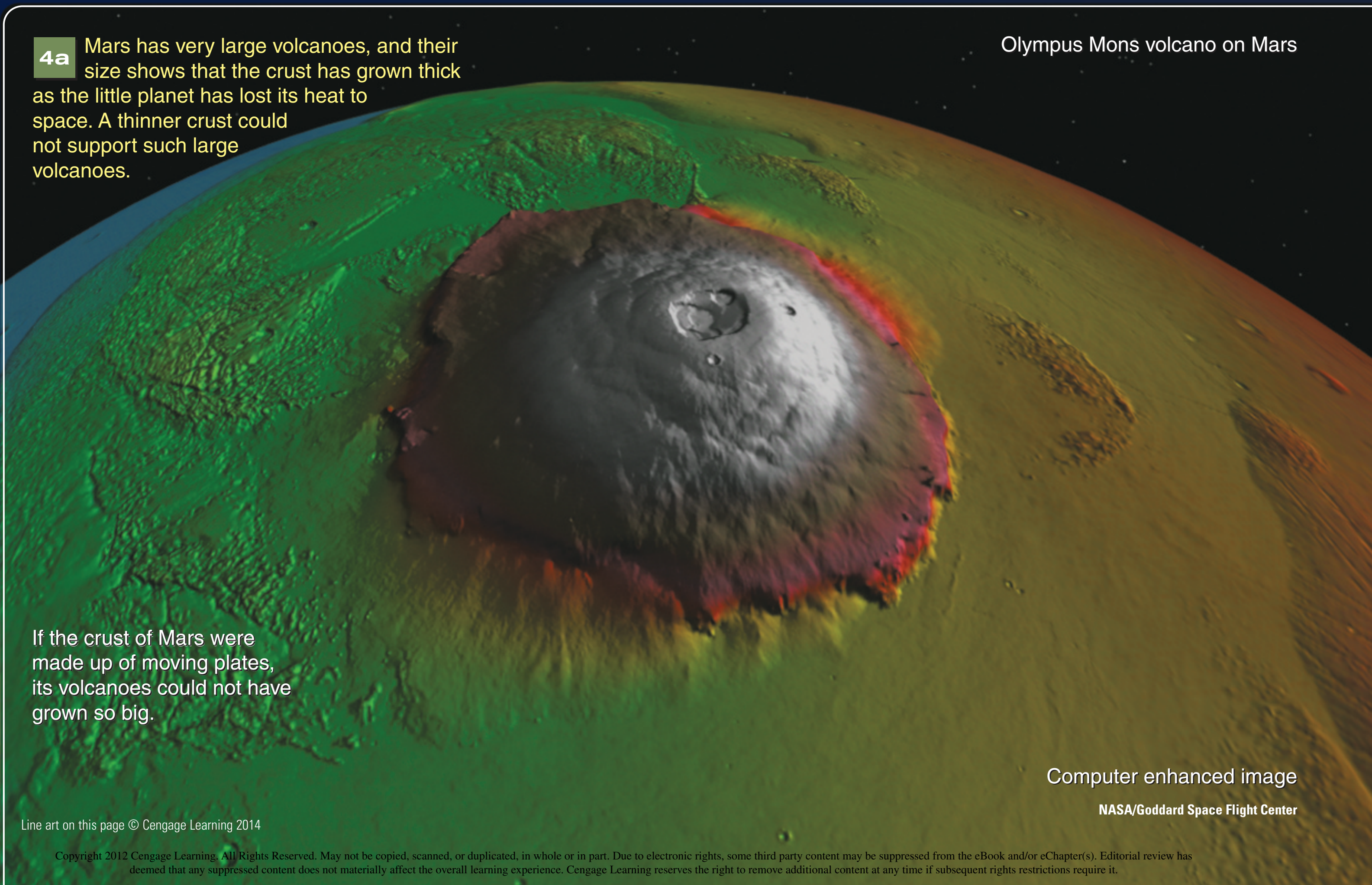
NASA/JPL/Malin Space Science Systems

4 Early in the history of Mars when its crust was thin, convection in the mantle could have pushed up volcanic regions such as Tharsis and Elysium, and limited plate motion could have produced Valles Marineris, but Mars cooled too fast, and its crust never broke into moving tectonic plates as did Earth's crust.

As its crust thickened, volcanism abated, and Mars lost the ability to replenish its waning atmosphere.

4a Mars has very large volcanoes, and their size shows that the crust has grown thick as the little planet has lost its heat to space. A thinner crust could not support such large volcanoes.

Olympus Mons volcano on Mars



If the crust of Mars were made up of moving plates, its volcanoes could not have grown so big.

Computer enhanced image

NASA/Goddard Space Flight Center

Line art on this page © Cengage Learning 2014

How Do We Know? 10-2

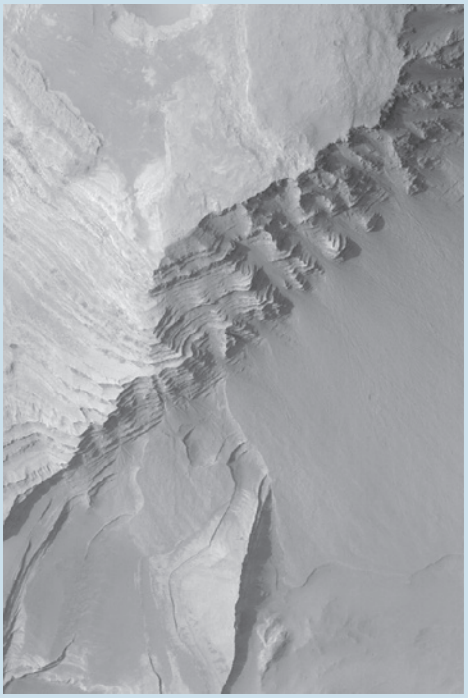
The Present Is the Key to the Past

How can we know what happened long ago if there weren't any witnesses? Geologists are fond of saying "The present is the key to the past." By that they mean that you can learn about the history of Earth by looking at the present condition of Earth's surface. The position and composition of various rock layers in the Grand Canyon, for example, tell you that the western United States was once at the floor of an ocean. This principle of geology is relevant today as you try to understand the history of other worlds such as Venus and Mars.

In the late 1700s, naturalists first recognized that the present gave them clues to the history of Earth. At the time that was astonishing because most people assumed that Earth had no history. That is, they assumed either that Earth had been created in its present state as described in the Old Testament or that Earth was eternal. In either case, people commonly assumed that the hills and mountains they saw around them had always existed more or less as they were. By the 1700s, naturalists began to see evidence that the hills and mountains were not eternal but were the result of past processes and were slowly changing. That gave rise to the idea that Earth had a history.

As those naturalists made the first attempts to thoughtfully and logically explain the nature of Earth by looking at the evidence, they were inventing modern geology as a way of understanding Earth. What Copernicus, Kepler, and Newton did for the heavens in the 1500s and 1600s, the first geologists did for Earth beginning in the 1700s. Of course, the invention of geology as the study of Earth led directly to the modern attempts to understand the geology of other worlds.

Geologists and astronomers share a common goal: They are attempting to reconstruct the past (How Do We Know? 8-2). Whether you study Earth, Venus, or Mars, you are looking at the present evidence and trying to reconstruct the past history of the planet by drawing on observations and logic to test each step in the story. How did Venus get to be covered with lava, and how did Mars lose its atmosphere? The final goal of planetary astronomy is to draw together all of the available evidence (the present) to tell the story (the past) of how the planet got to be the way it is. Those first geologists in the 1700s would be fascinated by the stories planetary astronomers tell today.



NASA/JPL/Malin Space Science Systems

Layers of rock in the Martian crater Terby hint at a time when the crater was filled with a lake.

(Figure 10-15b). It seems likely that Deimos experienced many collisions in its past, so its fractures may be hidden below the debris.

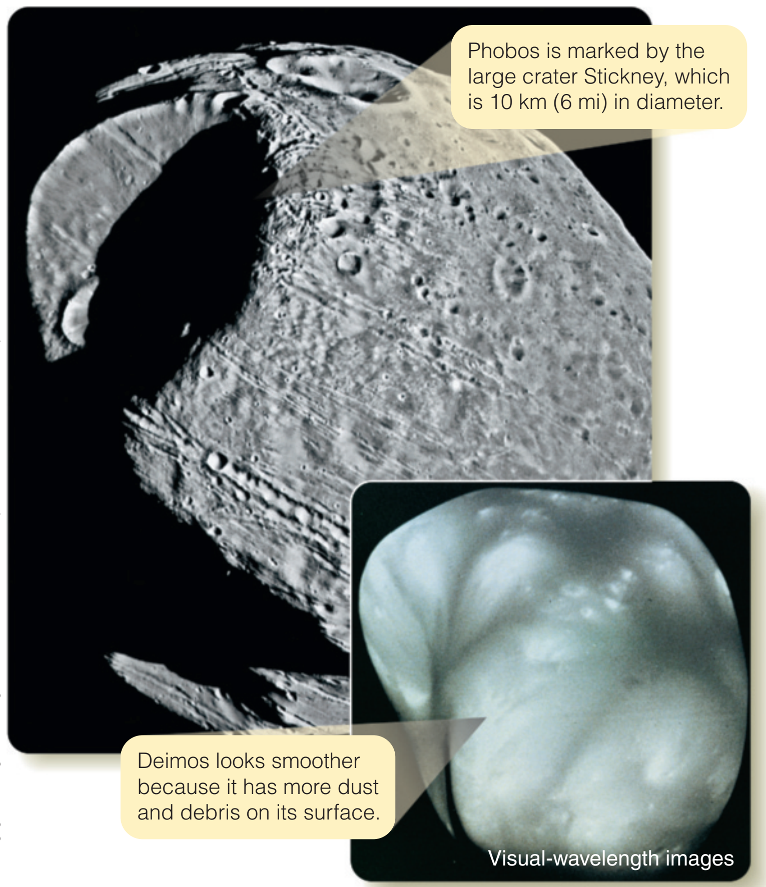
The debris on the surfaces of the moons raises an interesting question: How can the weak gravity of small bodies hold any fragments from meteorite impacts? The escape velocity on Phobos is only about 12 m/s (40 ft/s). An athletic astronaut who could jump 2 m (6 ft) high on Earth could jump 2.8 km (1.7 mi) on Phobos. Most of the fragments from an impact should escape, but the slowest particles could fall back in the weak gravity and accumulate on the surface. That cushioning layer in turn could help trap more debris.

Because Deimos is smaller than Phobos, its escape velocity is smaller, so it seems surprising that it has more debris on its surface. This may be related to Phobos's orbit close to Mars. The Martian gravity is almost strong enough to pull loose material off of Phobos's surface, so Phobos may be able to retain less of its cratering impact debris.

SCIENTIFIC ARGUMENT

Why would you be surprised to find volcanism on Phobos or Deimos? This argument hinges on the principle that the larger a world is, the more slowly it loses its internal heat. It is the flow of that heat from the interior through the surface into space that drives geologic activity such as volcanism and plate motion. A small world, like Earth's moon, cools quickly and remains geologically active for a shorter time than does a larger world like Earth. Phobos and Deimos are not just small; they are tiny. However they formed, any interior heat would have leaked away very quickly; with no energy flowing outward, there can be no volcanism.

Some futurists suggest that the first human missions to Mars will not land on the planet's surface but instead will build a colony on (or inside) Phobos or Deimos. These plans speculate that there may be water deep inside the moons that colonists could use. **What would happen to water released in the sunlight on the surface of such small worlds?**



© Cengage Learning 2014; images: Damon Simonelli and Joseph Veverka, Cornell University/NASA; Deimos: NASA

Figure 10-15

The moons of Mars are too small to pull themselves into spherical shape. The two moons, shown here to scale, were named for the mythical attendants of the god of war, Mars. Phobos was the god of fear, and Deimos was the god of dread.

What Are We? Comfortable

Many planets in the universe probably look like the moon and Mercury—small, airless, and cratered. Some are made of stone; and some, because they formed farther from their star, are made mostly of ices. If you randomly visited a planet anywhere in the universe, you would probably stand on a moonscape.

Earth is unusual but not rare. The Milky Way Galaxy contains over 100 billion stars, and over 100 billion galaxies are visible with existing telescopes. Most of those stars probably have planets, and although many planets may look like Earth's moon and Mercury, there also must be plenty of Earth-like worlds.

As you look around your planet, you should feel comfortable living on such a beautiful planet, but it was not always such a nice place. The craters on the moon and the moon rocks returned by astronauts show that the moon formed with a sea of magma. Mercury seems to have had a similar history, so Earth may have formed the same way. Its surface was once a seething ocean of liquid rock swathed in a hot, thick atmosphere, torn by eruptions of more rock, explosions of gas from the interior, and occasional impacts from space. The moon and Mercury assure you that that is the way Terrestrial planets begin. Earth has evolved to become your home world, but Mother Earth has had a violent past.

Study and Review

Summary

- Mercury is smaller than Earth but larger than Earth's moon. It has an old, heavily cratered surface and only a sporadic extremely thin atmosphere.
- Mercury has a much higher density than Earth's moon and must have a large metallic core. Mercury may have suffered a major impact that blasted away low-density crustal rock and left it with a metallic core that is large in proportion to the diameter of the planet.
- **Lobate scarps (p. 189)** are long curving cliffs formed by compression on Mercury when its large metallic core solidified and contracted.
- Although Venus is almost as large as Earth, it has a thick, cloudy atmosphere of carbon dioxide that hides the surface from sight. It can be studied by radar mapping.
- Venus's carbon dioxide atmosphere drives an intense greenhouse effect and makes that planet's surface hot enough to melt lead.
- Venus is slightly closer to the sun than Earth, too warm for liquid water oceans to dissolve carbon dioxide from the atmosphere easily, and warm enough to start evaporating its oceans, leading to a runaway greenhouse effect.
- **Composite volcanoes (p. 194)** are associated with plate tectonics subduction zones and plate boundaries on Earth, whereas **shield volcanoes (p. 194)**, found on Earth, Venus, and Mars, are caused by rising columns of magma called hot spots that break through the crust from below.
- Geologic activity on Venus is dominated not by plate tectonics but rather by volcanism and vertical tectonics, including **coronae (p. 192)**, large circular uplifted regions. Crater counts show that the entire surface of Venus has been covered over or replaced in the past half billion years.

- Mars is about half the size of Earth; it has a thin atmosphere and has lost much, but not all, of its internal heat.
- The loss of atmospheric gases depends on the size of a planet and its temperature. Mars is cold, but it is small and has a low escape velocity, and thus many of its lighter gases have leaked away.
- Some water may have leaked away from Mars as ultraviolet radiation from the sun broke it into hydrogen and oxygen, but some water is frozen in the polar caps and as **permafrost (p. 196)** in the soil.
- **Outflow channels (p. 201)** and **valley networks (p. 201)** visible from Mars orbiters seem to have been cut by sudden floods or by longer-term drainage, but water cannot exist as a liquid on Mars now because of its low temperature and low atmospheric pressure. Therefore, conditions on Mars must have once been different, allowing liquid water to flow on the surface.
- The southern hemisphere of Mars is old cratered terrain, but some large volcanoes lie in the north. The size of these volcanoes strongly suggests that the crust does not move horizontally.
- Some volcanism may still occur on Mars, but because the planet is small, it has cooled and is not very active geologically.
- Orbiters have found evidence of large amounts of water frozen below the surface.
- Robot rovers have found clear signs that the Martian climate was different in the past and that liquid water flowed over the surface in at least some places. The northern lowlands may even have held an ocean at one time.
- The two moons of Mars are probably captured asteroids. They are small, airless, and cratered. They lost their internal heat long ago.

Review Questions

1. Why doesn't Venus have as many craters as Earth's moon or Venus?
2. Why does Mercury have lobate scarps but Earth, Venus, and Mars do not?
3. What evidence indicates that plate tectonics does not occur on Venus? On Mars?
4. What evidence suggests that Venus has been resurfaced within the last half billion years?
5. Why are the atmospheres of Venus and Mars mostly carbon dioxide? Why is the atmosphere of Venus very dense but the atmosphere of Mars is very thin?
6. Why did Venus suffer from the runaway greenhouse effect but Earth has avoided it so far?
7. What evidence indicates that there has been liquid water on Mars?
8. Why do astronomers conclude that the crust on Mars must be thicker than Earth's crust?
9. What evidence indicates that the climate of Mars has cyclical changes?
10. **How Do We Know?** How can learning a hypothesis about the history of a planet improve your understanding of that planet?
11. **How Do We Know?** How is solving a crime that had no witnesses similar to figuring out what happened to Venus and Mars's oceans?

Discussion Questions

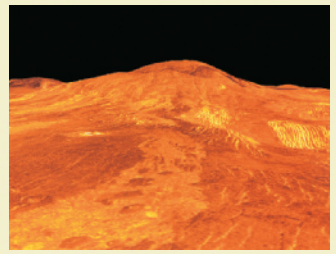
1. If liquid water is rare on the surface of planets, then most Terrestrial planets must have CO_2-rich atmospheres. Why?
2. If you visited a planet in another solar system and discovered that it was heavily cratered, but its small moon was nearly crater free, why would that be a surprise? Speculate about what might have happened to those objects.

Problems

1. Imagine that a spacecraft has landed on Mercury and is transmitting radio signals to Earth at a wavelength of 10.000 cm. When Mercury is seen from Earth in the evening sky, at its greatest angular distance east of the sun, it is moving toward Earth at its maximum possible relative speed of 47.9 km/s. To what wavelength must you tune your radio telescope to detect the signals? (*Hint:* Use the Doppler formula, Chapter 6.)
2. How long would it take radio signals to travel from Earth to Venus and back if Venus were at its nearest approach to Earth? Repeat the calculation for Venus at its farthest distance from Earth. (*Hint:* Useful data can be found in Appendix Table A-10.)
3. The smallest detail visible through Earth-based telescopes is about 1 second of arc in diameter. What size object would this represent on Mars at its closest approach to Earth? (*Hint:* Use the small-angle formula, Chapter 3.)
4. What is the maximum angular diameter of Phobos as seen from Earth? Useful data can be found in the chapter text and in Appendix Table A-10. (*Hint:* Use the small-angle formula, Chapter 3.)
5. Phobos orbits Mars at a distance of 9380 km from the center of the planet and has a period of 0.3189 days. Assume Phobos's orbit is circular. Calculate the mass of Mars. (*Hint:* Use the circular orbit velocity formula, Chapter 4. Remember to use units of meters, kilograms, and seconds.)

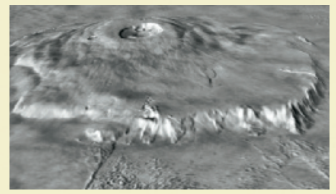
Learning to Look

1. Volcano Sif Mons on Venus is shown in this radar image. What kind of volcano is it, and why is it orange in this image? What color would the rock be if you could see it with your own eyes?



NASA/JPL

2. Olympus Mons on Mars is a very large volcano. In this image you can see multiple calderas superimposed at the top. What do those multiple calderas and the immense size of Olympus Mons indicate about the geology of Mars?



NASA

Great Debates

1. *Missions to Mars.* Life on Mars has been the subject of great debates, especially in the last century. The detection of life on another world would have revolutionary scientific, philosophical, and theological importance, but detecting such life with robotic probes is very difficult. All missions to Mars so far have been unmanned. However, several manned missions have been proposed and supported by past U.S. presidents like George W. Bush. Should manned or unmanned missions be sent to Mars? What are the benefits to society? What are the drawbacks?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

2. *Probes to Venus.* Venera 13 and other spacecraft that landed on Venus survived the high heat and pressures for only a few hours, but in that time they returned information about the geology and weather on a highly volcanic world with an intense greenhouse effect. Is sending unmanned spacecraft on deliberate kamikaze missions a wise use of taxpayer dollars? Why or why not?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

3. *Terraforming Terrestrials.* The year is 2100, and the world has become overrun with humans. However, we now have the technology to colonize the Terrestrial planets. You have been appointed to a committee to handle the overcrowding situation on Earth. You decide to sell real estate on Mercury, Venus, Earth's moon, and Mars. Three kinds of colonies must be created: (1) new correctional facilities for the government, (2) luxury and vacation homes for the wealthy, (3) middle-class housing, like the suburbs on Earth. Which Terrestrial planet is best suited for which colony and why?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

4. *Moon Base.* Imagine that some entrepreneurs in the United States decided to build a secret moon base. Using commercial space vehicles, they secretly resupplied the moon base with fresh scientists and equipment. For the last half-century, the secret project has been a success on the whole. It's time to let the world know about the findings. Should the scientific findings be made public, so all countries can share in the science and discoveries, or should the United States keep the most significant (or all) data for American scientists? What if the scenario changed, and the colonization of the moon was achieved by another nation? Would your answer change?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

5. *Penalties for Breaking Earth's Laws.* The year is 2100, and we now have the capability of commercially colonizing the moon. The time has come to invoke penalties for those who violate Earth's laws. You have been assigned to develop the criteria for which types of crimes involve imprisoning an Earthling to a correctional facility on the moon. Is a terrorist activity involving commercial flights to and from the Space Station a crime punishable by imprisonment to the moon's correctional facility? What minimum crime would necessitate imprisonment on the Moon?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

6. *Water on Mars.* Should we continue spending tax dollars going to Mars and searching for underground lakes of water? After all, Earth has underground lakes of liquid water in unlikely places like Antarctica. Why would we want to know if Mars has hidden water?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Lunar Craters
- Subduction and Volcanism
- Plate Motion Over a Hot Spot

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 5: Comparative Planetology

If you think about it, you might realize that *geology* is an odd word to apply to planets other than Earth. The Greek root word *geo* means "Earth," and *logy* is a version of *logos*, meaning "word." Thus, it seems *geology* should be words about Earth, full stop.

On the other hand, if the plan is to learn about other planets by comparing them to Earth and to learn about Earth by comparing our home planet to its siblings, then the field of study that has come to be known as "comparative planetology" is much the same as "planetary geology": words about Earth in the context of the other planets.

As you learned in this chapter, robot probes have examined the rocks and atmospheres of Venus and Mars while sitting or rolling on the surfaces of those planets. But those studies on location have explored only tiny fractions of the planetary surfaces. The surfaces of Venus and Mars together have five times the total area of Earth's continents, and even the long-lived Mars rovers *Spirit* and *Opportunity* traveled only a few miles from their landing sites.

The surfaces of Mercury, Venus, and Mars have been studied most thoroughly, albeit from distances of a few hundred kilometers, by orbiting spacecraft. Those probes have made thousands of photographs and radar images, measured heights and depths of surface features with laser altimeters, gathered surface reflectance spectra, counted gamma rays and neutrons, detected magnetic fields, and mapped planetary mass distributions using small anomalies in probe trajectories.

You learned in this and the preceding chapter that a few basic facts stand out from the wealth of details. The uncompressed average densities of the Terrestrial planets have a decreasing trend with distance from the sun, indicating a trend in the composition of the raw materials from which they each formed. All the Terrestrial planets plus Earth's moon Luna are differentiated, with relatively high-density material in their cores and lowest-density material at their surfaces. That fact indicates all of these objects must have been entirely liquid in the past, allowing materials of different densities to move and settle in their current locations.

The degree of surface "geological" activity runs in a neat sequence from the largest and most active of these five worlds, Earth, to the smallest and least active, Luna, consistent with the expected rates of cooling since their formation. Only Earth shows signs of planetwide plate tectonics, with horizontal crust movements producing most of the variety of our planet's landforms. The magnetic field strengths for the Terrestrial planets, offering another way to glimpse conditions in their cores, do not follow a pattern expected from their sizes and rotation rates and are somewhat of a mystery. All four planets plus Luna have been subjected to a continuous bombardment from asteroids and comets that was heaviest when the solar system was young. Some planetary surfaces show regions in which craters have been erased or covered over by subsequent activity.

Section 3 of Virtual Astronomy Lab 5, "Planetary Geology," lets you compare Mercury, Venus, Earth, Luna, and Mars, focused first on their respective rates of cooling from an original entirely molten state. That characteristic in turn explains many of their other similarities and differences. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

11

The Outer Solar System

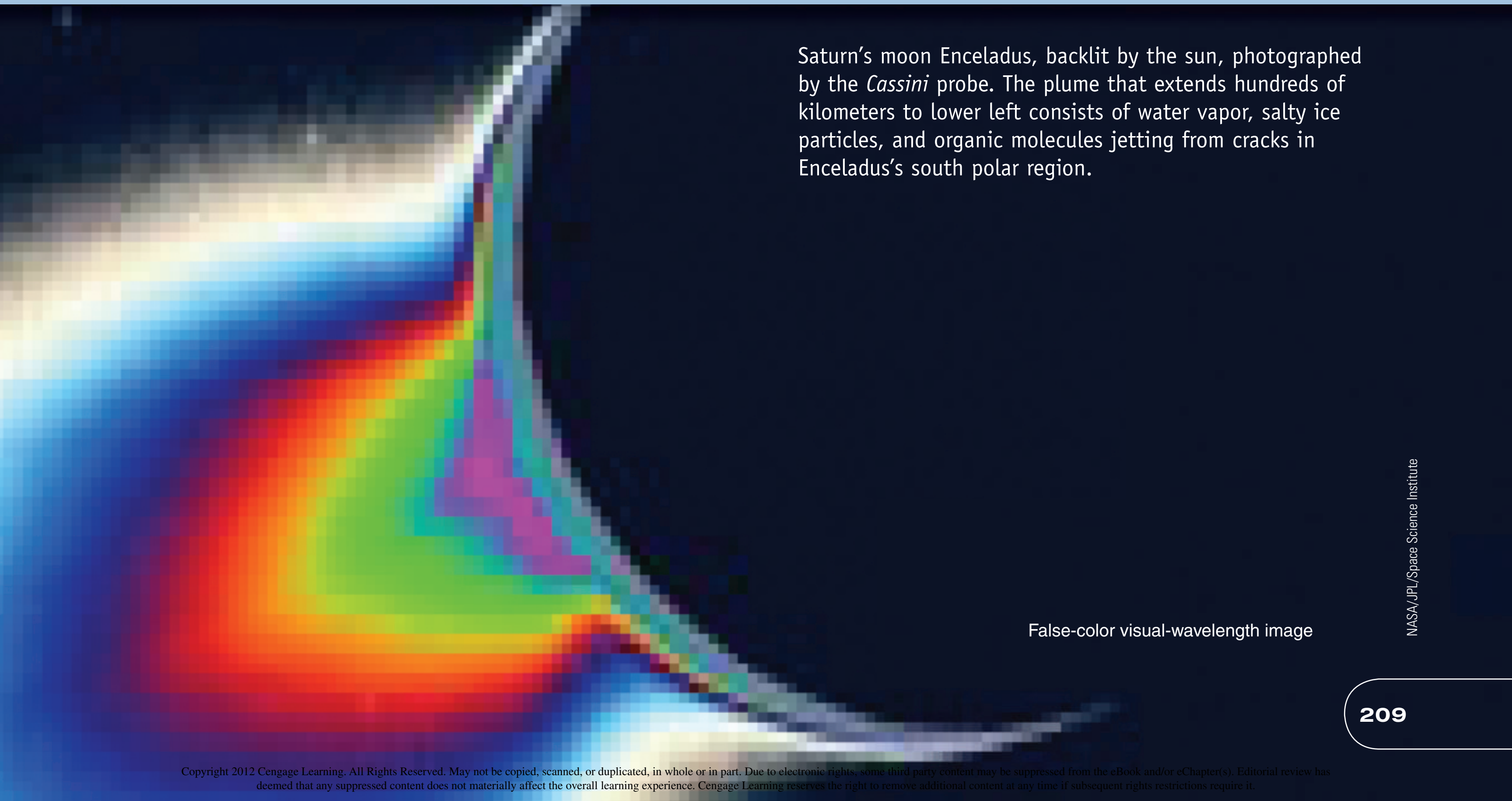
Guidepost

In the last three chapters, you watched our solar system form, then studied Earth and the other Terrestrial planets as they have evolved to the present day. In this chapter, the Jovian (Jupiter-like) planets will be a challenge to your imagination; they are such alien worlds, they would be unbelievable if you didn't have direct observational evidence to tell you what they are like. In contrast, Pluto and the other bodies in the Kuiper belt are not planets but intriguing remnants of the planet construction process.

As you explore, you will find answers to four important questions:

- **What are the properties of the Jovian planets?**
- **What is the evidence that some moons in the outer solar system have been geologically active?**
- **How are planetary rings formed and maintained?**
- **What do Pluto and the other Kuiper belt objects tell us about the formation of the solar system?**

The planets of the solar system formed from small bodies in the solar nebula, and the planets are still affected by the remnants of those populations of small objects. In the next chapter, you will adapt your understanding of comparative planetology to study the last bits of the solar nebula: meteorites, asteroids, and comets, and their continuing effects on Earth and the other planets.



Saturn's moon Enceladus, backlit by the sun, photographed by the *Cassini* probe. The plume that extends hundreds of kilometers to lower left consists of water vapor, salty ice particles, and organic molecules jetting from cracks in Enceladus's south polar region.

There wasn't a breath in that land of death . . .

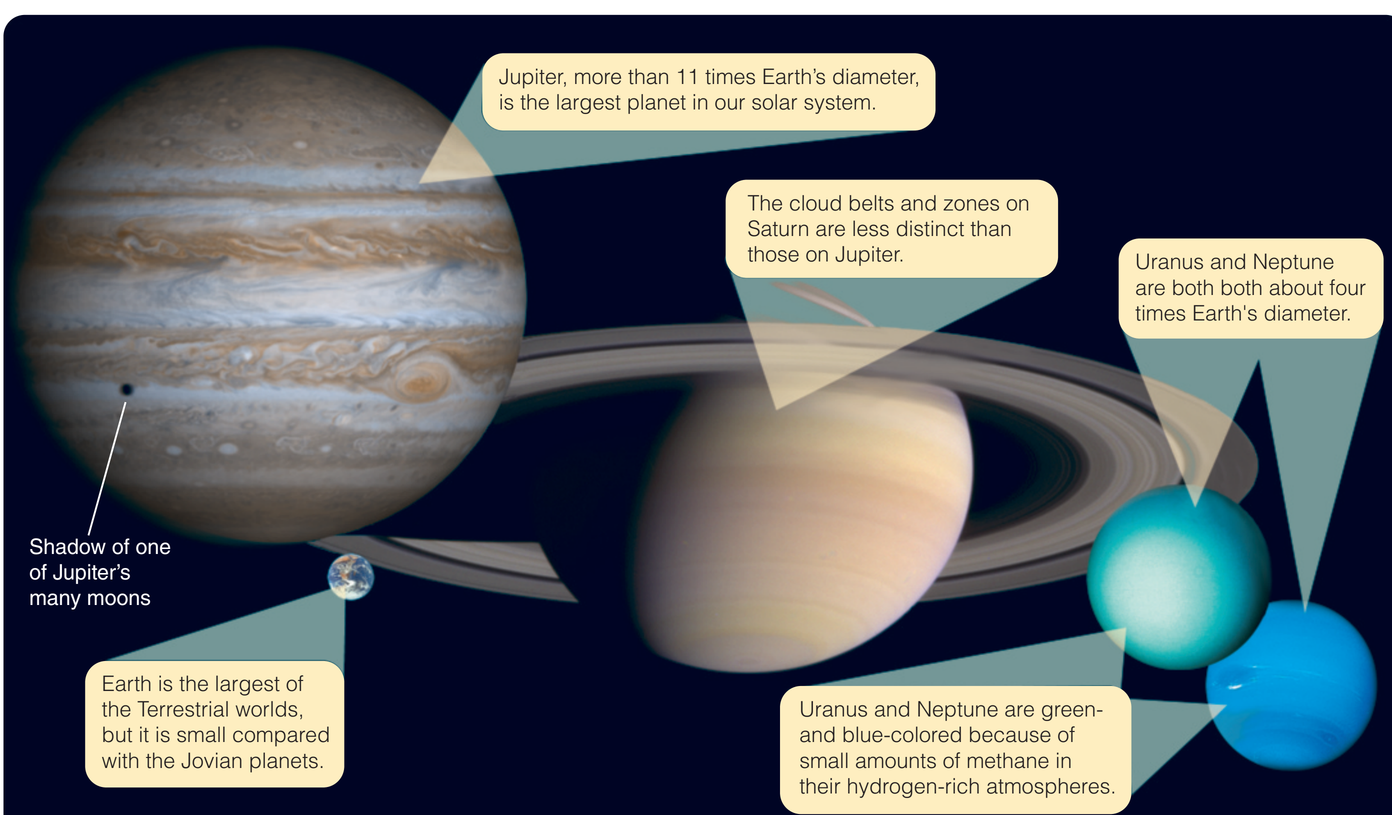
ROBERT SERVICE
THE CREMATION OF SAM MCGEE

THE SULFURIC acid clouds of Venus may seem totally alien to you, but compared with the planets of the outer solar system, Venus is almost like home. For example, the four Jovian planets have no solid surfaces.

The worlds of the outer solar system can be studied from Earth, but much of what astronomers know has been radioed back to Earth from space probes. The *Pioneer* and *Voyager* probes flew past the outer planets in the 1970s and 1980s, the *Galileo* probe orbited Jupiter in the late 1990s, and the *Cassini/Huygens* orbiter and probe arrived at Saturn in 2004. The *New Horizons* probe will pass Pluto in 2015 and then sail deeper into the Kuiper belt. Throughout this discussion, you will find images and data returned by these robotic explorers.

Figure 11-1

The principal worlds of the outer solar system are the four massive but low-density Jovian planets, each much larger than Earth.



© Cengage Learning 2014; Images: NASA/JPL/Space Science Institute/University of Arizona

11-1 A Travel Guide to the Outer Solar System

YOU ARE ABOUT TO VISIT worlds that are truly un-Earthly. This travel guide will warn you what to expect.

The Jovian Planets

The outermost planets in our solar system are Jupiter, Saturn, Uranus, and Neptune—the "Jovian planets," meaning they resemble Jupiter. In fact, they are each individuals with separate personalities. ■ Figure 11-1 compares the four outer worlds to each other and to Earth. One striking feature is their size. Jupiter is the largest of the Jovian worlds, over 11 times the diameter of Earth. Saturn is slightly smaller, and Uranus and Neptune are quite a bit smaller than Jupiter and Saturn, but still four times the size of Earth. Pluto, not pictured in the figure, is smaller than Earth's moon but was considered a planet from the time of its discovery in 1930 until 2006, when it was reclassified as a dwarf planet. You will learn about Pluto's characteristics, and the reasons for that decision, in this chapter.

The other feature you will notice immediately when you look at Figure 11-1 is Saturn's rings. They are bright and beautiful and composed of billions of ice particles. Jupiter, Uranus, and

Neptune also have rings, but they are not easily detected from Earth and are not visible in this figure. Nevertheless, as you visit these worlds you will be able to compare and contrast those four sets of planetary rings.

Atmospheres and Interiors

The four Jovian worlds have hydrogen-rich atmospheres filled with clouds. On Jupiter and Saturn, you can see that the clouds form stripes and bands that circle each planet, known as **belt-zone circulation.** You will find traces of these same types of features on Uranus and Neptune, but they are less distinct.

Models based on observations indicate that the atmospheres of the Jovian planets are not very deep; for example, Jupiter's atmosphere makes up only about one percent of its radius. Below their atmospheres Jupiter and Saturn are mostly liquid, so the conventional term for these planets, *gas giants,* should probably be changed to *liquid giants.* Uranus and Neptune are sometimes called *ice giants* because they contain abundant water in solid forms. Only near their centers do the Jovian planets have cores of dense material with the composition of rock and metal. None of the Jovian worlds has a definite solid surface on which you could walk.

Satellite Systems

You can't really land your spaceship on the Jovian worlds, but you might be able to land on one of their moons. All of the Jovian worlds have extensive satellite systems. Those moons can be classified into two groups: (1) the **regular satellites,** which tend to be large and orbit relatively close to their parent planet, with low inclinations to the planet's equator, moving in the **prograde** direction along with most of the objects in the solar system; versus (2) the **irregular satellites,** which tend to be smaller than the regular satellites, sometimes have retrograde and/or highly inclined orbits, and are generally far from their parent planet. Astronomers have evidence that the regular satellites formed approximately where they are now as the planets formed but that the irregular satellites are mostly, if not all, captured objects.

Some of the Jovian moons are geologically active now, while others show signs of past activity. Of course, geological activity depends on heat flow from the interior, so you might ponder what could be heating the insides of these small objects.

SCIENTIFIC ARGUMENT

Why do you expect the outer planets to be low-density worlds?
This should be a familiar argument. In Chapter 8 you discovered that the inner planets could not incorporate ice when they formed because it was too hot at their locations near the sun. By contrast, in the outer solar nebula, water vapor could freeze to form ice particles. The icy particles accumulated rapidly into Jovian protoplanets with densities lower than that of the rocky Terrestrial planets and asteroids. Consequently, the Jovian planets grew massive enough to pull in even lower-density hydrogen and helium gas directly from the nebula by gravitational collapse. The ices and gas made the outer planets low-density worlds.

Now you can expand your argument. **Why do you expect the outer planets to have high-density cores?**

11-2 Jupiter

Jupiter is the largest and most massive of the Jovian planets, containing 71 percent of all the planetary matter in the entire solar system. Just as you used Earth, the largest of the Terrestrial planets, as the basis for comparison with the others, you can examine Jupiter in detail as a standard in your comparative study of the other Jovian planets.

The Interior

Although Jupiter is very large, it is only 1.3 times denser than water (**Celestial Profile 7**, page 219). For comparison, Earth is more than 5.5 times denser than water. As you have already learned, the density of a planet is an important clue about the average composition of the planet's interior. Jupiter's shape also gives information about its interior. Jupiter and the other Jovian planets are all slightly flattened. A world with a large rocky core and mantle would not be flattened much by rotation, but an all-liquid planet would flatten significantly. Thus Jupiter's **oblateness,** the fraction by which its equatorial diameter exceeds its polar diameter, combined with its average density, helps astronomers calculate what its insides are like.

Models show that the interior of Jupiter is mostly liquid hydrogen. However, if you jumped into Jupiter carrying a kayak, expecting an ocean, you would be disappointed. The base of the atmosphere is so hot and the pressure is so high that there is no sudden boundary between liquid and gas. As you fell deeper and deeper through the atmosphere, you would find the gas density increasing around you until you were sinking through a liquid, but you would never splash into a distinct liquid surface.

Under very high pressure, liquid hydrogen becomes **liquid metallic hydrogen**—a material that is a very good conductor of electricity. Model calculations indicate that most of Jupiter's interior is composed of this material. That large mass of conducting liquid, stirred by convection currents and spun by the planet's rapid rotation, drives the dynamo effect and generates a powerful magnetic field. Jupiter's field is over ten times stronger than Earth's.

A planet's magnetic field deflects the solar wind and dominates a volume of space around the planet called the **magnetosphere.** Jupiter's magnetosphere is 100 times larger than Earth's (■ Figure 11-2a). If you could see it in the sky, it would be more than five times larger than the full moon. Just as in the case of Earth (look back to Chapter 7), interactions

between Jupiter's magnetic field and the solar wind generate powerful electric currents that flow around the planet's magnetic poles. These are visible at ultraviolet wavelengths as rings of aurora lights that are larger in diameter than Earth (Figure 11-2b).

The strong magnetic field around Jupiter traps charged particles from the solar wind in radiation belts a billion times more intense than the Van Allen belts that surround Earth. The spacecraft that have flown through these regions received over 4000 times the radiation that would have been lethal for a human.

At Jupiter's center, a so-called rocky core contains heavier elements, such as iron, nickel, silicon, and so on. With a temperature four times hotter than the surface of the sun and a pressure 50 million times Earth's sea-level atmospheric pressure, this material is unlike any rock on Earth. The term *rocky core* refers to the chemical composition, not to the mechanical properties of the material.

Careful measurements of the heat flowing out of Jupiter reveal that it emits about twice as much energy as it absorbs from the sun. This energy appears to be heat left over from the formation of the planet. In Chapter 8 you saw that Jupiter should have grown very hot when it formed, and some of that heat remains inside, slowly leaking into space. Look back to Figure 5-17b on page 92; this infrared image of Jupiter from the *SOFIA* airborne telescope shows a white stripe where the planet's internal heat is prominent.

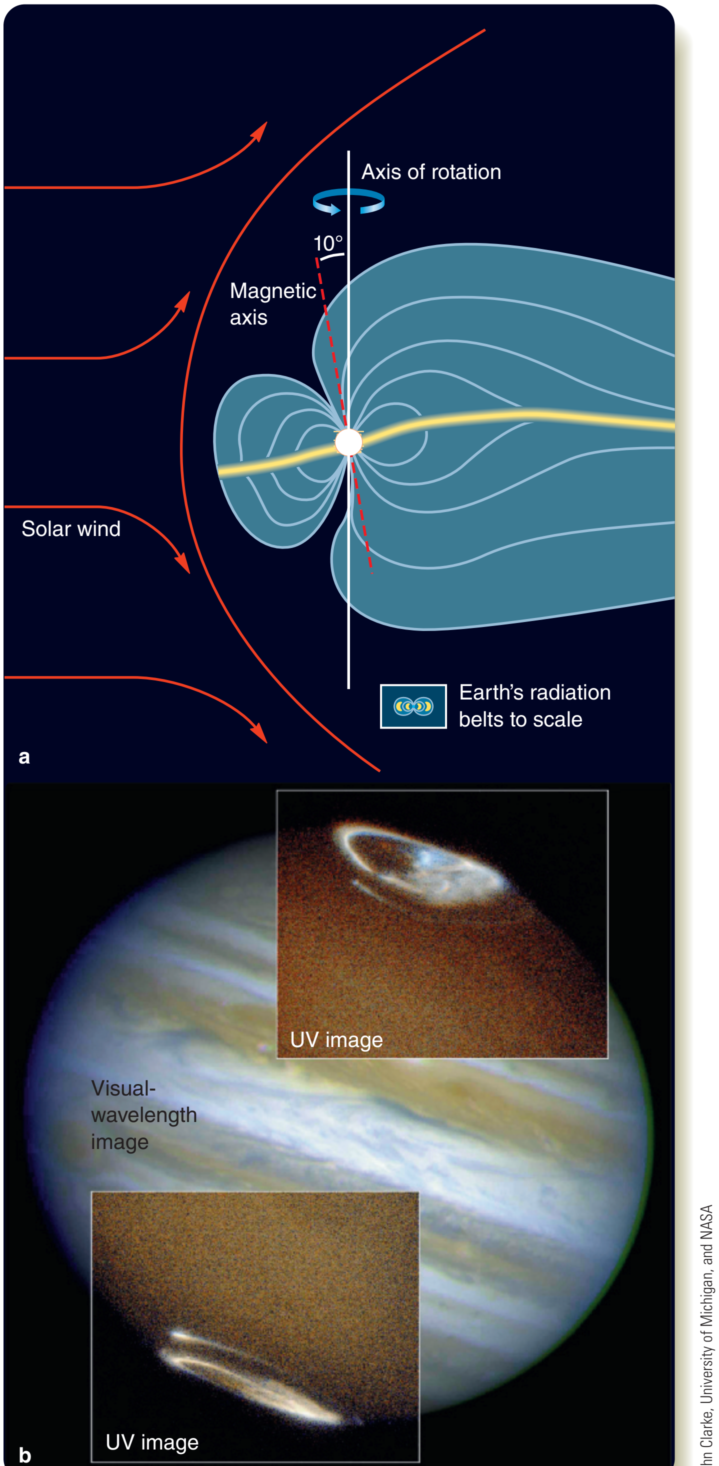


Figure 11-2

(a) Jupiter's large conducting core and rapid rotation create a powerful magnetic field that holds back the solar wind and dominates the magnetosphere, colored blue in the diagram. High-energy particles trapped in the magnetic field form giant radiation belts. (b) Auroras on Jupiter are confined to rings around the north magnetic pole and the south magnetic pole, as shown in these ultraviolet images. Earth's auroras follow the same pattern. The small comet-shaped spots are caused by powerful electrical currents flowing from Jupiter's moon Io.

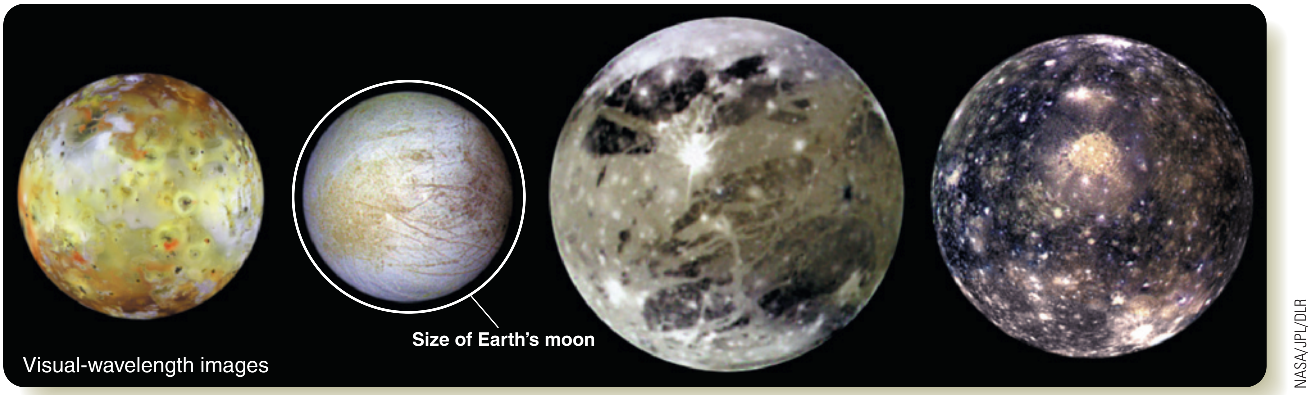
Jupiter's Complex Atmosphere

It is a **Common Misconception** that Jupiter is a ball of gases. In fact, as you have learned, Jupiter is almost entirely a liquid planet. Its atmosphere is only a thin outer skin of turbulent gases and clouds. The processes you will find there are repeated in slightly different ways on the other Jovian worlds.

Study **Jupiter's Atmosphere** on pages 214–215 and notice four important ideas:

1. The atmosphere is hydrogen rich, and the clouds are confined to a shallow layer.
2. The cloud layers are located at certain temperatures within the atmosphere where ammonia (NH_3), ammonium hydrosulfide ($(NH_4)SH$), and water (H_2O), respectively, can condense.
3. The belt-zone circulation pattern of colored cloud bands circling the planet like stripes on a child's ball is related to the high- and low-pressure areas found in Earth's atmosphere.
4. The large circular or oval spots seen in Jupiter's clouds are circulating storms that can remain stable for decades or even centuries.

Photos of Jupiter lead to the **Common Misconception** that the clouds are at the top of the atmosphere. Notice that the atmosphere of transparent hydrogen and helium extends high above the cloud tops. What you see in photos is only the cloud layers.



■ **Figure 11-3**

Color-enhanced visual-wavelength images of the Galilean moons of Jupiter, from left to right: Io, Europa, Ganymede, and Callisto.

Jupiter's Moons

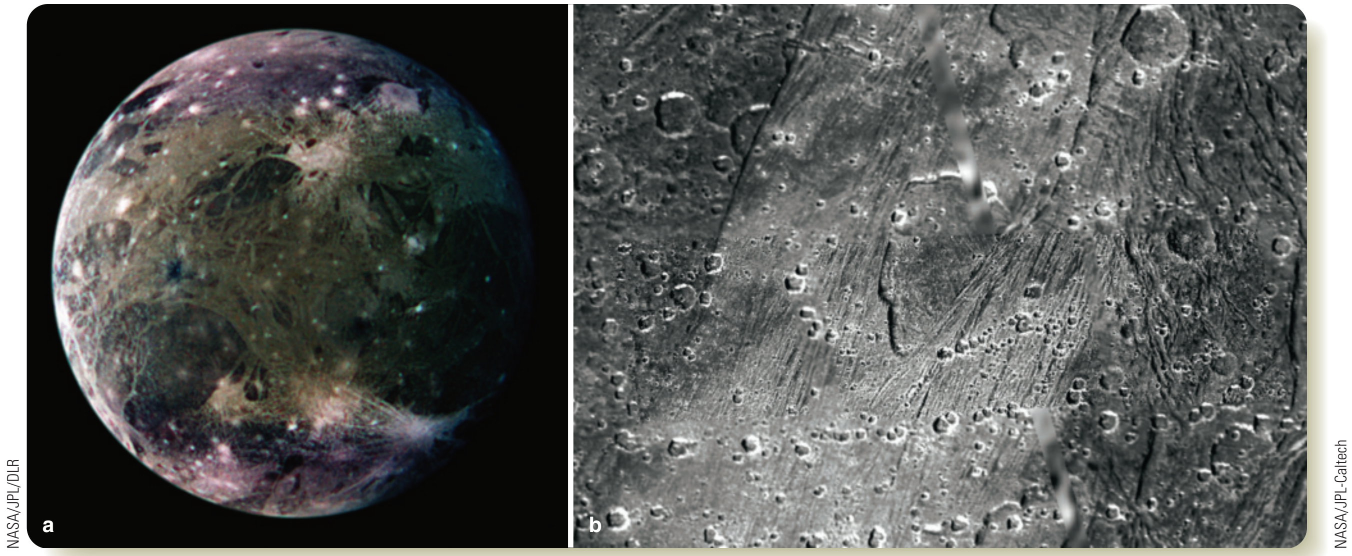
Jupiter has four large moons and at least 60 smaller moons. Larger telescopes and modern techniques are continually finding more small moons orbiting all the Jovian planets, and each of the Jovian planets probably has many more small undiscovered moons. Some of the small moons are probably captured asteroids. In contrast, the four largest moons of Jupiter (■ Figure 11-3), called the **Galilean moons** after their discoverer, Galileo, are clearly related to each other and almost certainly formed with Jupiter.

The outermost Galilean moons, Ganymede and Callisto, are about the size of Mercury, one and a half times the size of Earth's moon. In fact, Ganymede is the largest moon in the solar system. Ganymede and Callisto have low densities of only 1.9 and 1.8 g/cm^3, respectively, meaning they must consist roughly of half rock and half ice. Observations of their gravitational fields by the *Galileo* spacecraft reveal that Ganymede has a rocky or metallic core and lower-density icy exterior, so it is differentiated. Callisto is apparently only partly differentiated. Both moons interact with Jupiter's magnetic field in ways that show they probably have mineral-rich layers of liquid water 100 km or more below their icy crusts.

Callisto's surface and most of Ganymede's surface appear old because they are heavily cratered and very dark (■ Figure 11-4a). The continuous blast of meteoroids evaporates surface ice, leaving behind embedded minerals to form a dark skin like the grimy crust on an old snowbank. Thus, icy surfaces get darker with age. More recent impacts dig up cleaner ice and leave bright craters, as you can see on Callisto in Figure 11-3. Ganymede has some younger, brighter **grooved terrain** believed to be systems of faults in the

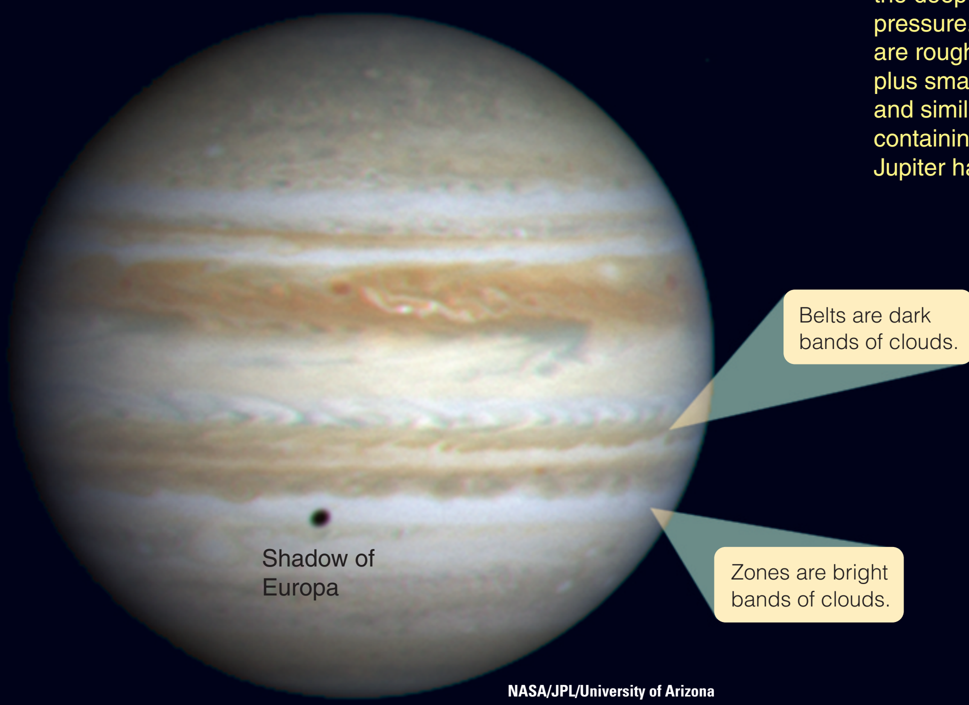
■ **Figure 11-4**

(a) This color-enhanced visual-wavelength image of Ganymede shows the frosty poles at top and bottom, the old dark terrain, and the brighter grooved terrain. (b) A band of bright terrain on Ganymede runs from lower left to upper right, and a collapsed area, a possible caldera, lies at the center in this visual-wavelength image. Calderas form where subsurface liquid has drained away, and the bright areas contain other features associated with flooding by water.



Jupiter's Atmosphere

1 Humans will probably never visit Jupiter's atmosphere. Its cloud layers are deathly cold, and the deeper layers that are warmer have a crushingly high pressure. There is no free oxygen to breathe; the gases are roughly three-quarters hydrogen and a quarter helium, plus small amounts of water vapor, methane, ammonia, and similar molecules. Traces of sulfur and molecules containing sulfur probably make it smell bad. Of course, Jupiter has no surface, so there isn't even a place to stand.

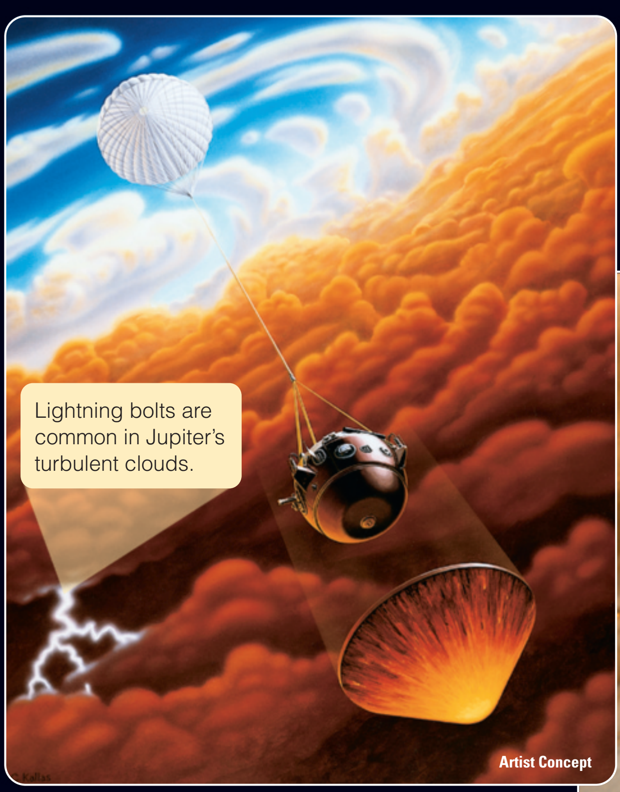


NASA/JPL/University of Arizona

Jupiter's moon Europa

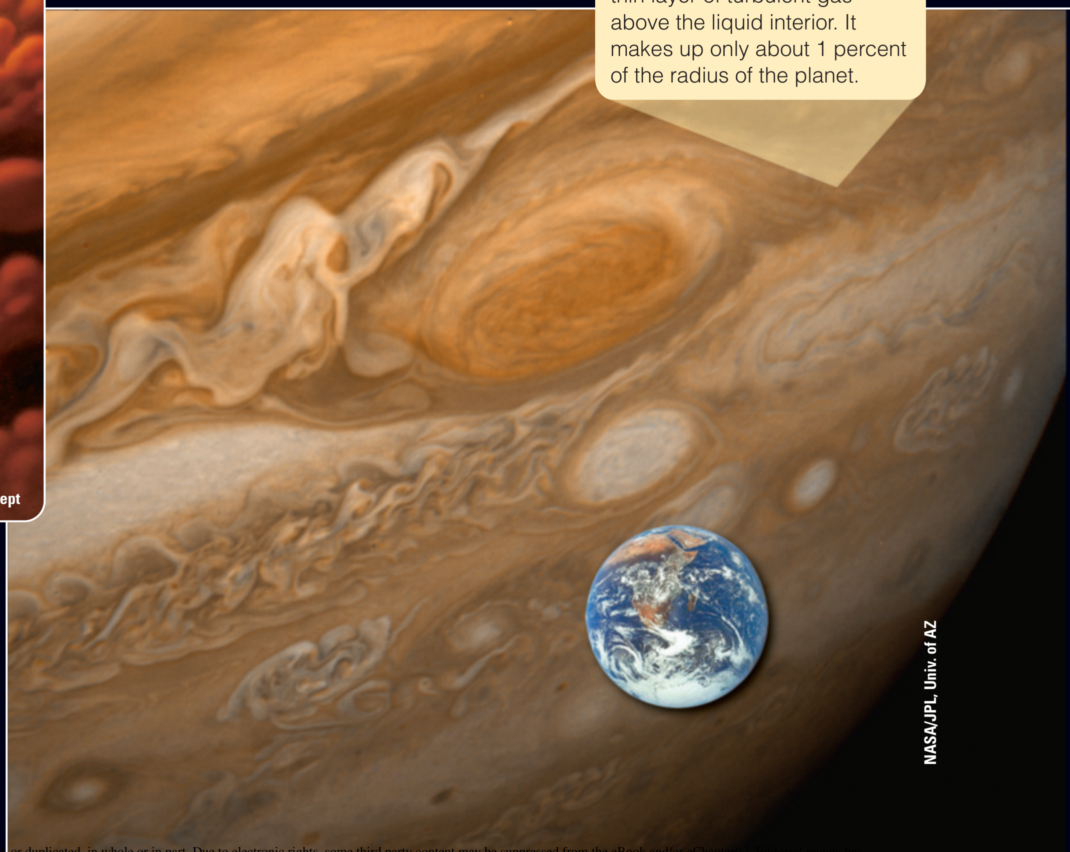

NASA/JPL/University of Arizona

1a The only spacecraft to enter Jupiter's atmosphere was the *Galileo* probe. Released from the *Galileo* spacecraft, the probe entered Jupiter's atmosphere in 1995. It parachuted through the upper atmosphere of clear hydrogen, dropped its heat shield, and then descended through layers of increasing pressure in Jupiter's stormy atmosphere until it was finally crushed.



Hughes Aircraft Company

Jupiter's atmosphere is a very thin layer of turbulent gas above the liquid interior. It makes up only about 1 percent of the radius of the planet.

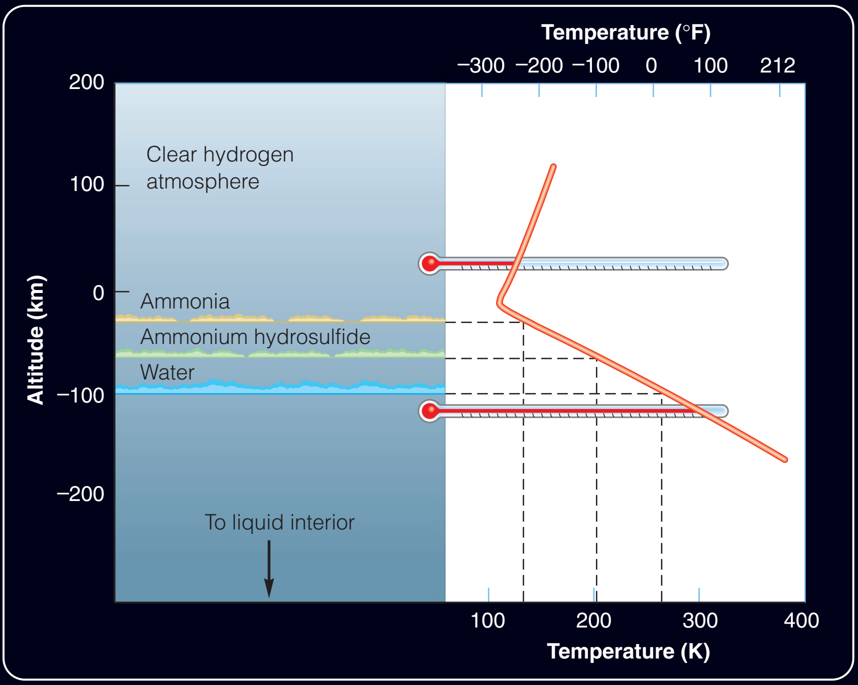

NASA/JPL, Univ. of AZ

The Great Red Spot at right is a giant circulating storm in one of the southern zones. It has lasted at least 350 years since astronomers first noticed it after the invention of the telescope. Smaller spots are also circulating storms.

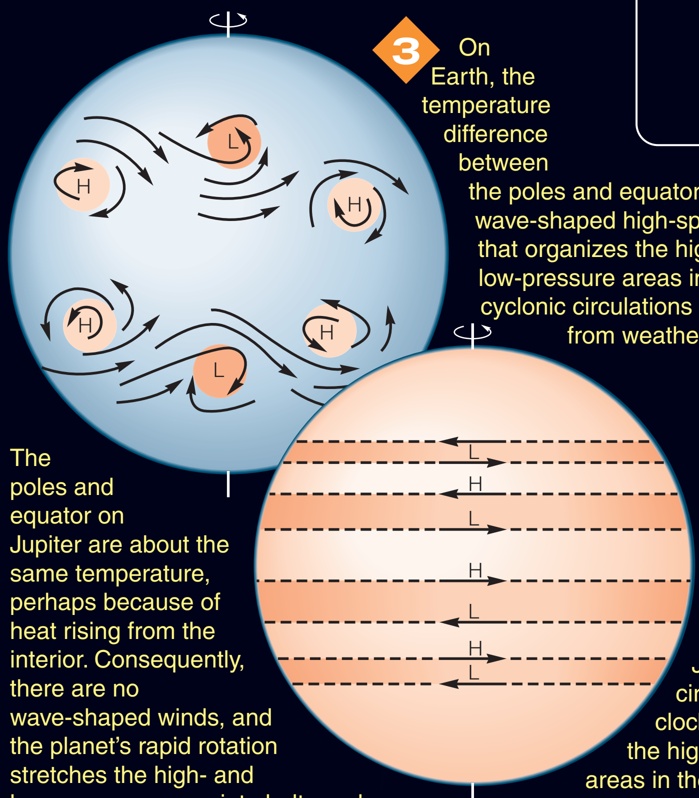
2 The visible clouds on Jupiter are composed of ammonia crystals, but models predict that deeper layers of clouds contain ammonium hydrosulfide crystals, and deeper still lies a cloud layer of water droplets. These compounds are normally white, so planetary scientists think the colors arise from small amounts of other molecules formed in reactions powered by lightning or sunlight.

If you could put thermometers in Jupiter's atmosphere at different levels, you would discover that the temperature rises below the uppermost clouds.

Far below the clouds, the temperature and pressure climb so high the gaseous atmosphere merges gradually with the liquid hydrogen interior and there is no surface.

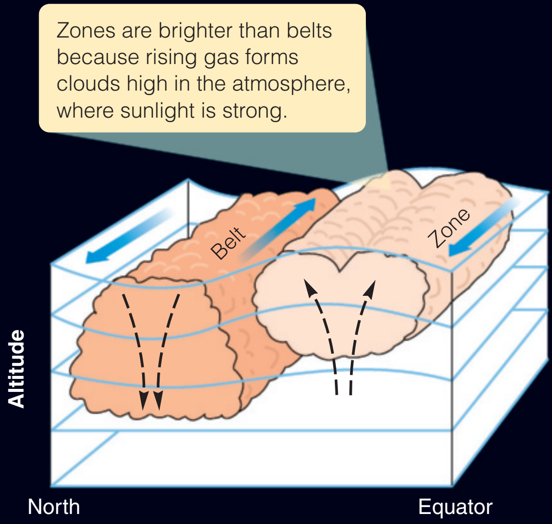


3 On Earth, the temperature difference between the poles and equator drives a wave-shaped high-speed wind that organizes the high- and low-pressure areas into cyclonic circulations familiar from weather maps.



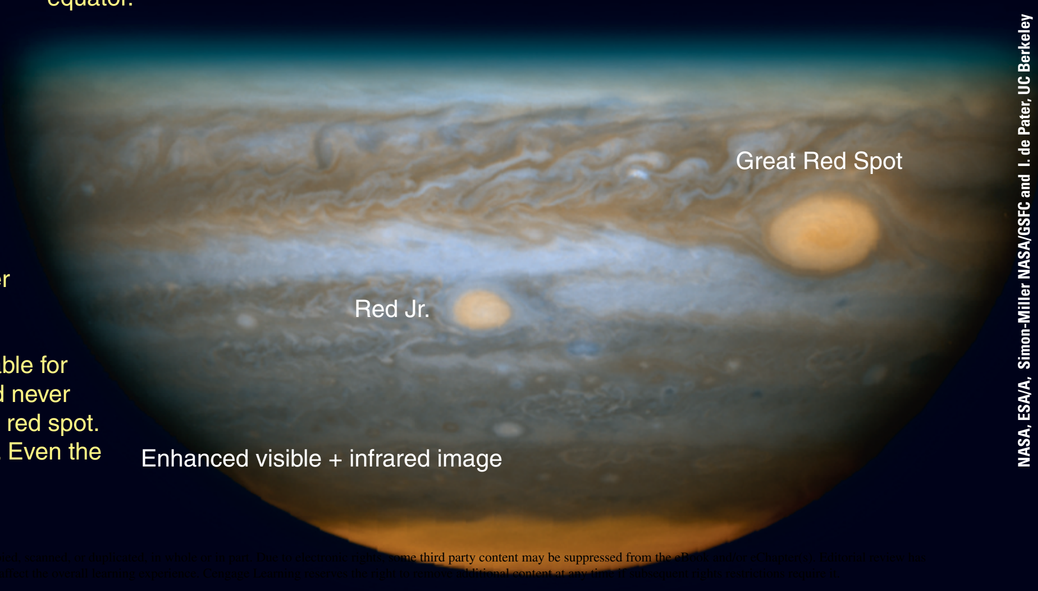
The poles and equator on Jupiter are about the same temperature, perhaps because of heat rising from the interior. Consequently, there are no wave-shaped winds, and the planet's rapid rotation stretches the high- and low-pressure areas into belts and zones that circle the planet.

On both Earth and Jupiter, winds circulate clockwise around the high-pressure areas in the northern hemisphere and counterclockwise south of the equator.

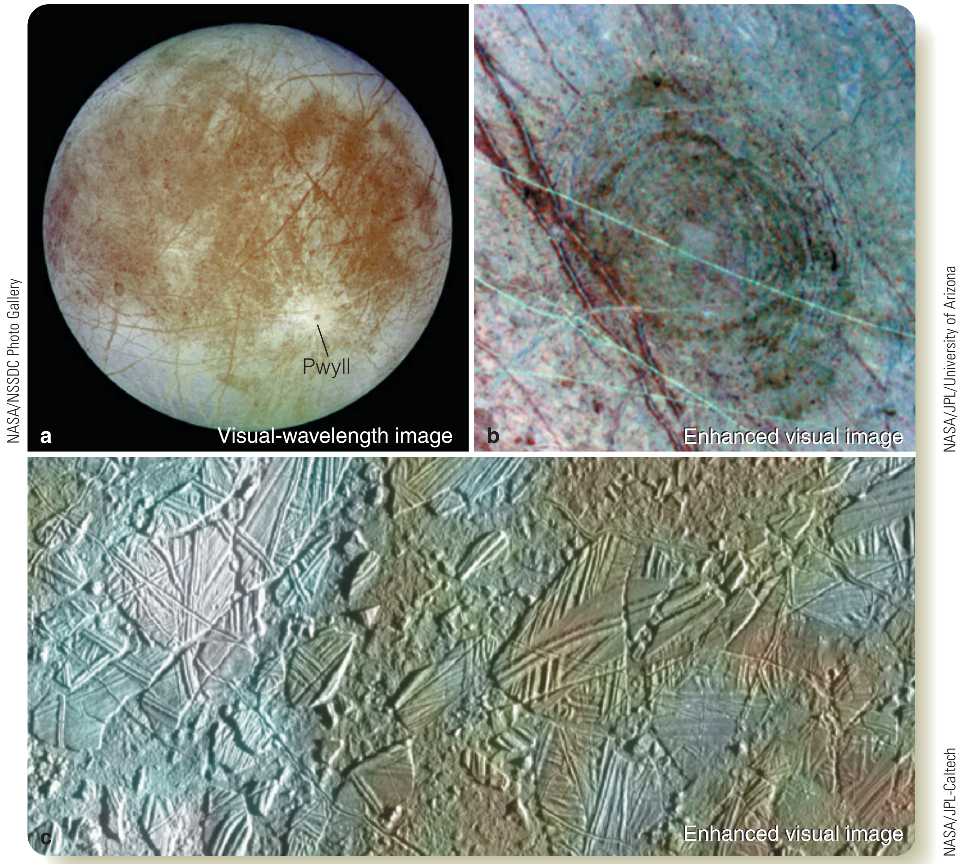


4 Three circulating storms visible as white ovals since the 1930s merged in 1998 to form a single white oval. In 2006, the storm intensified and turned red like the Great Red Spot, so it was dubbed "Red Jr." The reason for the red color is unknown, but it may show that the storm is bringing material up from lower in the atmosphere.

Storms in Jupiter's atmosphere may be stable for decades or centuries, but astronomers had never before witnessed the appearance of a new red spot. It may eventually vanish or develop further. Even the Great Red Spot may someday vanish.



NASA, ESA/A. Simon-Miller NASA/GSFC and I. de Pater, UC Berkeley



■ **Figure 11-5**

(a) The icy surface of Europa is shown here in natural color. Many faults are visible on its surface, but very few craters. The bright crater is Pwyll, a young impact feature. (b) This circular bull's-eye is the remains of a crater 140 km (85 mi) in diameter. Notice the younger cracks and faults that cross the older impact feature. (c) Like icebergs on an arctic ocean, blocks of crust on Europa appear to have floated apart and rotated. The blue icy surface is stained brown by mineral-rich water venting from below the crust. White areas are ejecta from the impact that formed Pwyll crater.

brittle crust. Some sets of grooves overlap other sets of grooves, suggesting extended episodes of geological activity (Figure 11-4b).

The density of the next moon inward, Europa, is 3.0 g/cm^3, which means that Europa is mostly rock with a thin icy crust. The visible surface is very clean ice, contains very few craters, has long cracks in the icy crust, and includes complicated terrain that resembles blocks of ice in Earth's Arctic Ocean (■ Figure 11-5). The pattern of mountainlike folds on its surface suggests that the icy crust breaks as the moon is flexed by tides (look back to Chapter 4). Europa's gravitational influence on the *Galileo* spacecraft reveals that a liquid-water ocean perhaps 200 km deep lies below the 10- to 100-km-thick crust. The lack of craters tells you that Europa is an active world where craters are quickly erased.

Images from spacecraft reveal that Io, the innermost of the four Galilean moons, has over 100 volcanic vents on its surface (■ Figure 11-6). The active volcanoes throw sulfur-rich gas and ash high above the surface. That ash falls back to bury the surface at a rate of a few millimeters a year. This explains why you see no impact craters on Io—they are covered up as fast as they form. Io's density is 3.6 g/cm^3, showing that it is composed of rock and metal. Its gravitational influence on the passing *Galileo* spacecraft revealed that it is differentiated into a large metallic core, a rocky mantle, and a low-density crust.

The activity you see in the Galilean moons must be driven by energy flowing outward, yet these objects are too small to have remained hot from the time of their formation. Io's volcanism seems to be driven by **tidal heating.** Io follows a slightly elliptical orbit caused by its interactions with the other moons. As Io's distance from Jupiter varies, the planet's gravitational field flexes the moon with varying tidal force, and the resulting friction heats Io's interior. That heat flowing outward causes the volcanism. Europa is not as active as Io, but it also must have a heat source, presumably tidal heating. Ganymede is no longer active, but when it was younger it must have had internal heat to break the crust and produce the grooved terrain.

In fact, those three moons are linked together in **orbital resonances.** Io orbits Jupiter four times while Europa orbits twice and Ganymede orbits once. These resonances keep all three orbits slightly elliptical and drive tidal heating that makes the moons active now, or made them active in the past. Distant Callisto is not caught in this orbital resonance and appears never to have been strongly active. Given the lack of surface activity, it is surprising that there is evidence of a subsurface ocean within Callisto as well as in Europa and Ganymede.

Jupiter's Rings

Astronomers have known for centuries that Saturn has rings, but Jupiter's ring was not discovered until 1979, when the *Voyager 1* spacecraft sent back photos. Less than 1 percent as bright as Saturn's icy rings, Jupiter's rings are very dark and reddish, showing that the material is rocky rather than icy.

Astronomers can also conclude that the ring particles are mostly microscopic. Photos of the ring show that it is very bright when illuminated from behind (■ Figure 11-7a)—called **forward scattering.** Large particles do not scatter light forward: A ring filled with basketball-size particles would look dark when illuminated from behind. Forward scattering tells you that Jupiter's rings are made of tiny grains with diameters approximately equal to the wavelengths of light, less than a millionth of a meter, about the size of particles in cigarette smoke.

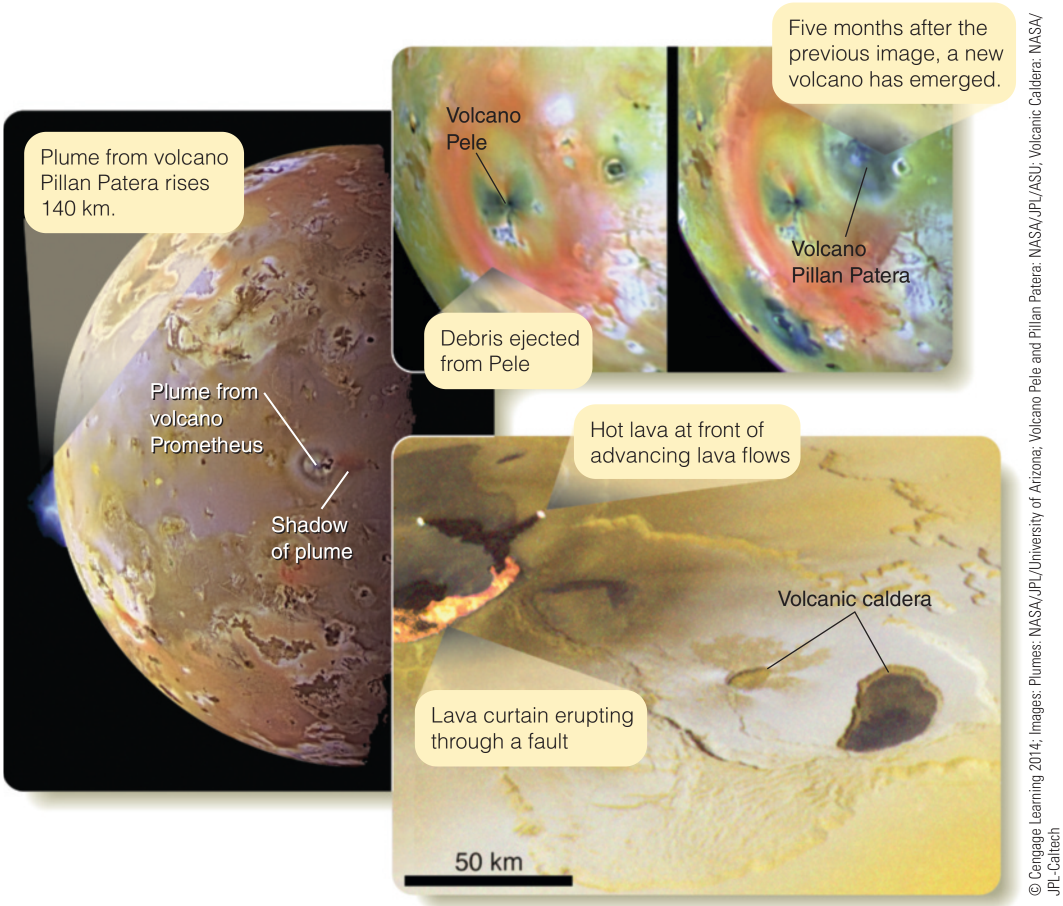


Figure 11-6

These enhanced-color images of volcanic features on Io were produced by combining visual and near-infrared images and digitally enhancing the colors. To human eyes, most of Io would look pale yellow and light orange.

The rings orbit inside the **Roche limit,** the distance from a planet within which a moon cannot hold itself together by its own gravity. If a moon were to come inside the Roche limit, tidal forces would overcome the moon's gravity and pull the moon apart. Also, raw material for a moon cannot coalesce inside the Roche limit. The Roche limit is about 2.4 times the planet's radius, depending somewhat on the relative densities of the planet and the orbiting material. The rings of Saturn, Uranus, and Neptune also lie inside the respective Roche limits for each planet.

Now you can understand Jupiter's dusty rings, If a dust speck gets knocked loose from a larger rock inside the Roche limit, the rock's gravity cannot hold the dust speck. For that same reason, the billions of dust specks in the rings can't pull themselves together to make a moon because of tidal forces inside the Roche limit.

You can be sure that Jupiter's ring particles are not old. The pressure of sunlight and the planet's powerful magnetic field can quickly alter the orbits of the particles. Images show faint ring material extending down toward the cloud tops, evidently dust specks spiraling toward Jupiter. Dust is also lost from the ring as electromagnetic effects force it out of the central plane to form a low-density halo above and below the ring (Figure 11-7b).

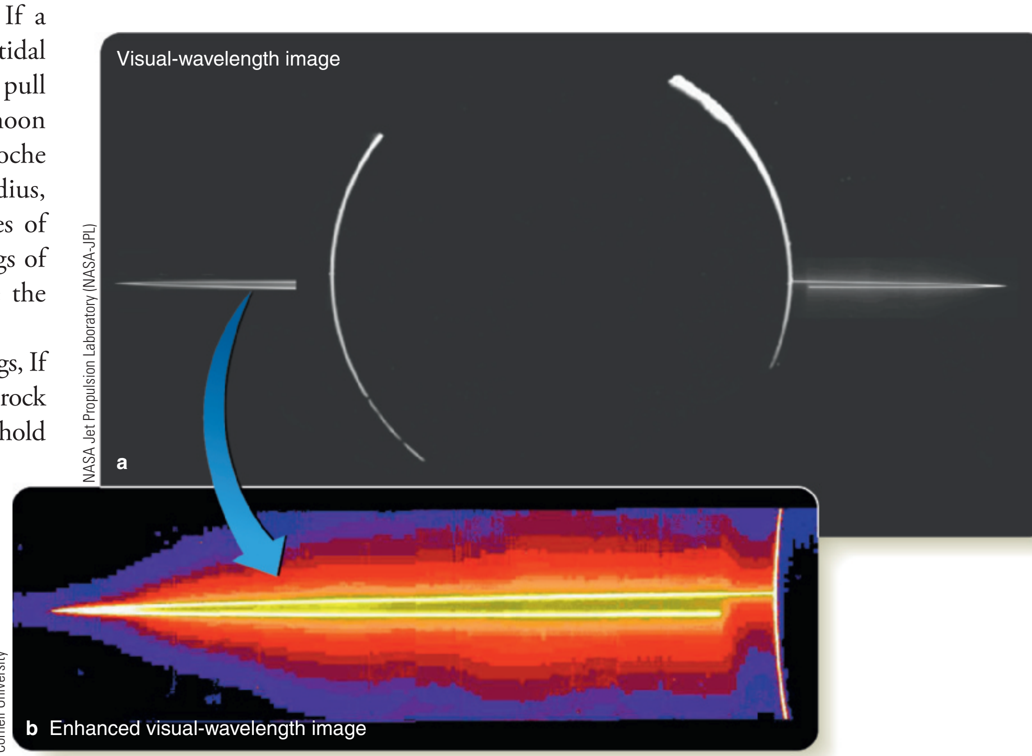


Figure 11-7

(a) The main ring of Jupiter, illuminated from behind, glows brightly in this visual-wavelength image made by the *Galileo* spacecraft located within Jupiter's shadow. (b) Digital enhancement and false color reveal the halo of ring particles that extends above and below the main ring. The halo is just visible in panel (a).

Another reason the ring particles can't be old is that the intense radiation around Jupiter tends to grind the particles down to nothing in a century or so. Therefore, the ring you see today can't be material left over in its current situation since the formation of Jupiter. Instead, the ring must be continuously resupplied with new dust. Observations made by the *Galileo* spacecraft provide evidence that the source of the ring material is meteoroids eroding the small moons Adrastea, Metis, Amalthea, and Thebe that orbit near or within the rings.

The rings around Saturn, Uranus, and Neptune are also known to be short lived, so they also must be resupplied by new material, probably eroded from nearby moons. Aside from supplying the Jovian planets' rings with particles, moons also act to confine the rings, keep them from spreading outward, and alter their shapes. You will explore these processes in detail when you study the rings of the other planets later in this chapter.

A History of Jupiter

Can you put all of the evidence together and tell the story of Jupiter? Creating such a logical argument of evidence and hypotheses is the ultimate goal of planetary astronomy.

Jupiter formed far enough from the sun to incorporate large numbers of icy planetesimals, and it must have grown rapidly. Once it was about 15 times more massive than Earth, it could grow by gravitational collapse (see Chapter 8), capturing gas directly from the solar nebula. Thus, it grew rich in hydrogen and helium from the solar nebula. Its present composition resembles the composition of the sun and the solar nebula. Jupiter's gravity is strong enough to hold onto all its gases, even hydrogen (look back to Figure 10-8).

The large family of moons may be mostly captured asteroids, and Jupiter may still encounter a wandering asteroid or comet now and then. Some of these are deflected, and some, like comet Shoemaker-Levy 9 that struck Jupiter in 1994, actually fall into the planet (see Chapter 12). Dust blasted off the inner moons by meteoroid impacts settles into the equatorial plane to form Jupiter's ring.

The four Galilean moons are large and seem to have formed like a mini-solar system in a disk of gas and dust around the forming planet. The innermost Galilean moon, Io, is densest, and the densities of the others decrease as you move away from Jupiter, similar to the way the densities of the planets decrease with distance from the sun. Perhaps the inner moons incorporated less ice because they formed closer to the heat of the growing planet. You can recognize that tidal heating also has been important, and the intense heating of the inner moons could have driven off much of their ices. Thus, a combination of two processes may be responsible for the compositions of the Galilean moons.

SCIENTIFIC ARGUMENT

Why is Jupiter so large?

You can analyze this question by constructing a logical argument that relates the formation of Jupiter to the solar nebula theory. Jupiter grew so rapidly from icy planetesimals in the outer solar nebula that it eventually became large enough to be able to continue growing by gravitational collapse. By the time the solar nebula cleared away and ended planet building, Jupiter had captured large amounts of hydrogen and helium. The Terrestrial planets are made from a small fraction of the elements present in the solar nebula, but the Jovian planets incorporated abundant hydrogen and helium and were able to become very massive.

Now create a new argument. **How is the activity on Io and Europa powered?**

11-3 Saturn

Saturn is most famous for its beautiful rings, easily visible through the telescopes of modern amateur astronomers. Large Earth-based telescopes have explored the planet for many decades. The first close-up views came when two *Voyager* probes flew past Saturn in 1980 and 1981. The *Cassini* spacecraft went into orbit around Saturn in 2004 and began an extended exploration of the planet, its rings, and its moons (**How Do We Know? 11-1**).

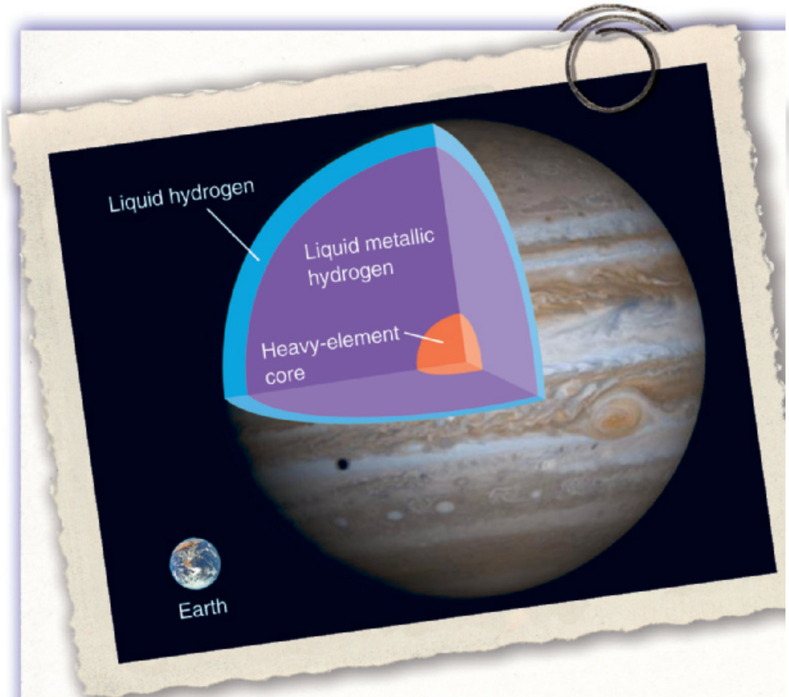
Saturn the Planet

As you can see in Figure 11-1, Saturn shows only faint belt–zone circulation, but *Voyager, Cassini,* and *Hubble Space Telescope* images confirm that belts and zones are present and that the associated winds are up to three times faster than the winds on Jupiter. Belts and zones on Saturn are less visible than on Jupiter because they occur deeper in Saturn's colder atmosphere, below a layer of methane haze (■ Figure 11-8).

Saturn is less dense than water (it would float!), suggesting that it is, like Jupiter, rich in hydrogen and helium. Saturn is the most oblate of the planets, and this is evidence that its interior is mostly liquid with only a small core of heavy elements (**Celestial Profile 8**, page 219). Because its internal pressure is lower, Saturn has less liquid metallic hydrogen than Jupiter. Perhaps that is why Saturn's magnetic field is 20 times weaker than Jupiter's. Like Jupiter, Saturn radiates more energy than it receives from the sun, and models predict that it, too, has a very hot interior.

Saturn's Moons

Saturn has nearly 50 known moons, many of which are small and all of which contain mixtures of ice and rock. Many are probably captured objects.



© Cengage Learning 2014; image: NASA/JPL/University of Arizona

Jupiter is mostly a liquid hydrogen planet with a small core of heavy elements that is not much bigger than Earth.

Celestial Profile 7: Jupiter

Motion:

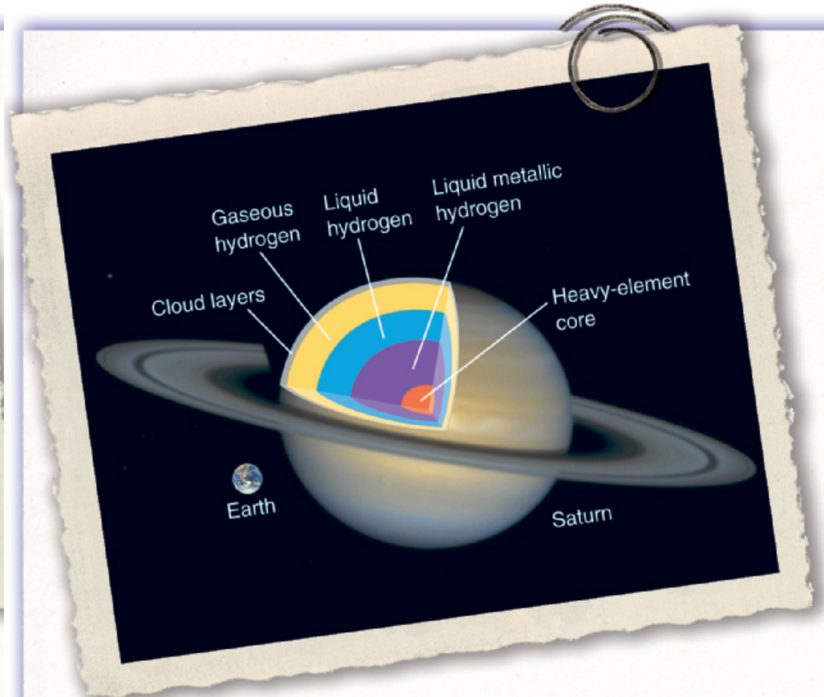
Average distance from the sun	5.20 AU (7.79 × 10^8 km)
Eccentricity of orbit	0.048
Inclination of orbit to ecliptic	1.3°
Orbital period	11.9 y
Period of rotation	9.92 h
Inclination of equator to orbit	3.1°

Characteristics:

Equatorial diameter	1.43 × 10^5 km (11.2 $D_\oplus$)
Mass	1.90 × 10^{27} kg (318 $M_\oplus$)
Average density	1.33 g/cm^3
Gravity at cloud tops	2.5 Earth gravities
Escape velocity	59.5 km/s (5.3 $V_\oplus$)
Temperature at cloud tops	145°K (−200°F)
Albedo	0.34
Oblateness	0.065

Personality Point:

Jupiter is named for the Roman king of the gods (the Greek Zeus), and it is the largest planet in our solar system. It can be very bright in the night sky, and its cloud belts and four largest moons can be seen through even a small telescope. Its moons are even visible with a good pair of binoculars mounted on a tripod.



© Cengage Learning 2014; image: NASA/STScI

Density, oblateness, and gravity measurements made by planetary probes allow planetary astronomers to model Saturn's interior.

Celestial Profile 8: Saturn

Motion:

Average distance from the sun	9.58 AU (1.43 × 10^9 km)
Eccentricity of orbit	0.056
Inclination of orbit to ecliptic	2.5°
Orbital period	29.5 y
Period of rotation	10.57 h
Inclination of equator to orbit	26.7°

Characteristics:

Equatorial diameter	1.21 × 10^5 km (9.45 $D_\oplus$)
Mass	5.68 × 10^{26} kg (95.2 $M_\oplus$)
Average density	0.69 g/cm^3
Gravity at cloud tops	1.1 Earth gravities
Escape velocity	35.5 km/s (3.2 $V_\oplus$)
Temperature at cloud tops	95°K (−290°F)
Albedo	0.34
Oblateness	0.098

Personality Point:

The Greek god Cronus was forced to flee when his son Zeus took power. Cronus went to Italy where the Romans called him Saturn, protector of the sowing of seed. He was celebrated in a weeklong Saturnalia festival at the time of the winter solstice in late December. Early Christians took over the holiday to celebrate Christmas.

Funding for Basic Research

Who pays for science? Searching out scientific knowledge can be expensive, and that raises the question of funding. Some science has direct applications, and industry supports such research. For example, pharmaceutical companies have large budgets for scientific research leading to the creation of new drugs. But some basic science is of no immediate practical value, with no obvious commercial applications. Who pays the bill?

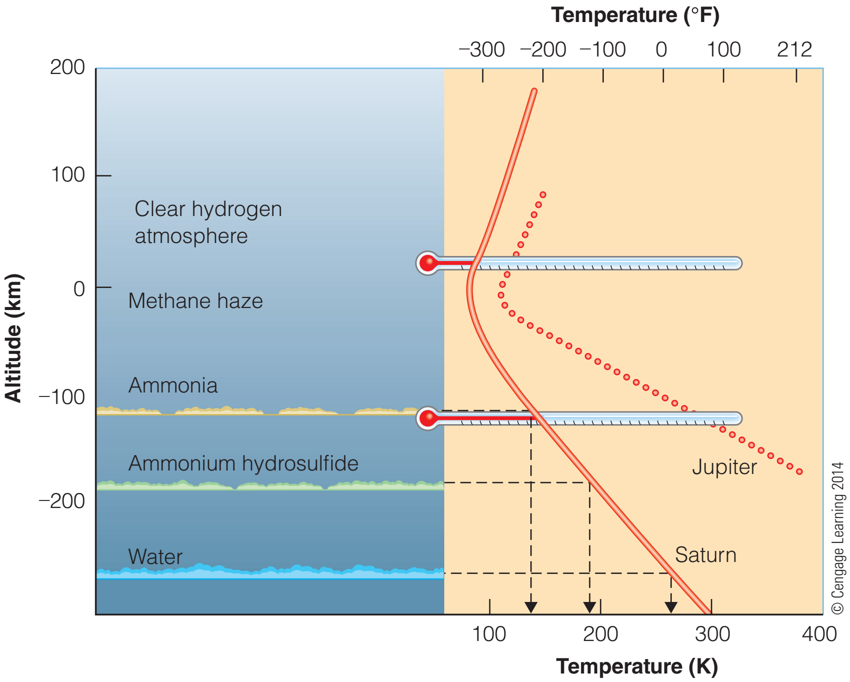
Should you consider a career as an industrial paleontologist? A paleontologist is a scientist who studies ancient life-forms by examining fossils of plant and animal remains. Except for the rare Hollywood producer about to release a dinosaur movie, corporations can't make a profit from the discovery of a new dinosaur. The practical-minded stockholders of a company will not approve major investments in such research. Consequently, digging up dinosaurs, like astronomy, is not well funded by industry.

It falls to government institutions and private foundations to pay the bill for this kind of research. The Keck Foundation built two giant telescopes on Mauna Kea in Hawai'i with no expectation of financial return, and the National Science Foundation has funded thousands of astronomy research projects for the benefit of society.

The discovery of a new dinosaur or a new galaxy is of no great financial value, but such scientific knowledge is not worthless. Its value lies in what it tells us about the world we live in. Such scientific research enriches our lives by helping us understand what we are. Ultimately, funding basic scientific research is a public responsibility that society must balance against other needs. There isn't anyone else to pick up the tab.

Exploring other worlds is valuable; it helps us humans understand ourselves and our place in the universe. Yet, sending the Cassini *spacecraft to Saturn costs each U.S. citizen only 56¢ per year over the life of the project.*

The largest of Saturn's moons is Titan, a bit larger than the planet Mercury. Measurements of the effect of its gravitational field on probes passing nearby indicate that it has a rocky core under a thick mantle of ices. Titan is so cold that its gas molecules do not travel fast enough to escape. It has an atmosphere composed mostly of nitrogen with traces of argon and methane. The ultraviolet component of sunlight converts some of the methane into complex carbon-rich molecules that collect into small particles, filling the atmosphere with orange smog (■ Figure 11-9a). Some of those particles evidently settle downward and coat parts of the surface with what has been described as organic "goo," meaning it is composed of carbon-rich molecules and is probably semiliquid.

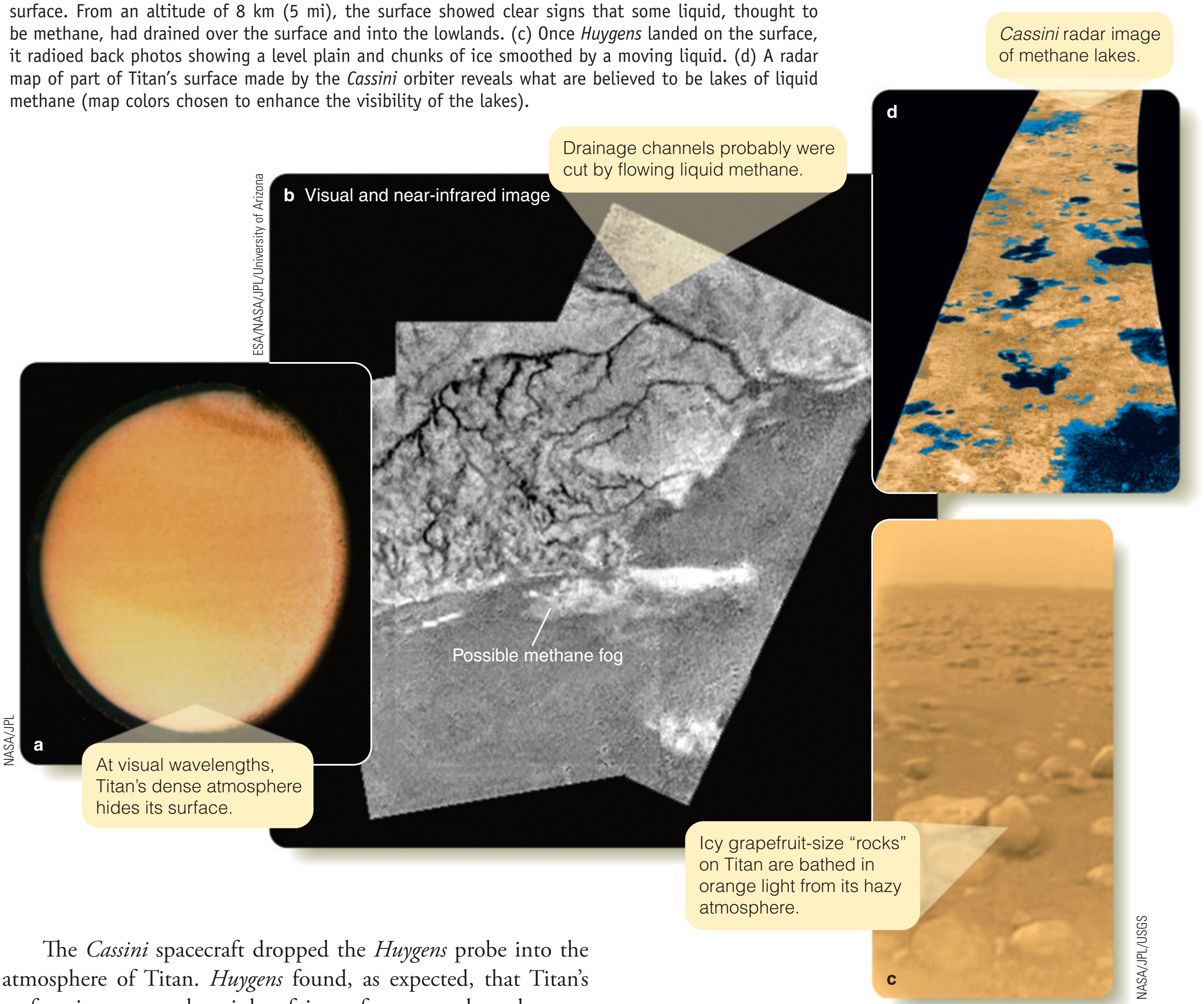


■ Figure 11-8

Because Saturn is farther from the sun, its atmosphere is colder (solid curve) than Jupiter's (dotted curve). The cloud layers on Saturn form at the same temperature as do the cloud layers on Jupiter, but that puts them deeper in Saturn's hazy atmosphere, where they are not as easy to see from the outside as Jupiter's clouds. (See page 215.)

Figure 11-9

(a) Saturn's largest moon, Titan, is surrounded by opaque orange clouds of organic particles. (b) The Huygens probe descended by parachute through the hazy atmosphere of Titan and photographed the surface. From an altitude of 8 km (5 mi), the surface showed clear signs that some liquid, thought to be methane, had drained over the surface and into the lowlands. (c) Once *Huygens* landed on the surface, it radioed back photos showing a level plain and chunks of ice smoothed by a moving liquid. (d) A radar map of part of Titan's surface made by the *Cassini* orbiter reveals what are believed to be lakes of liquid methane (map colors chosen to enhance the visibility of the lakes).

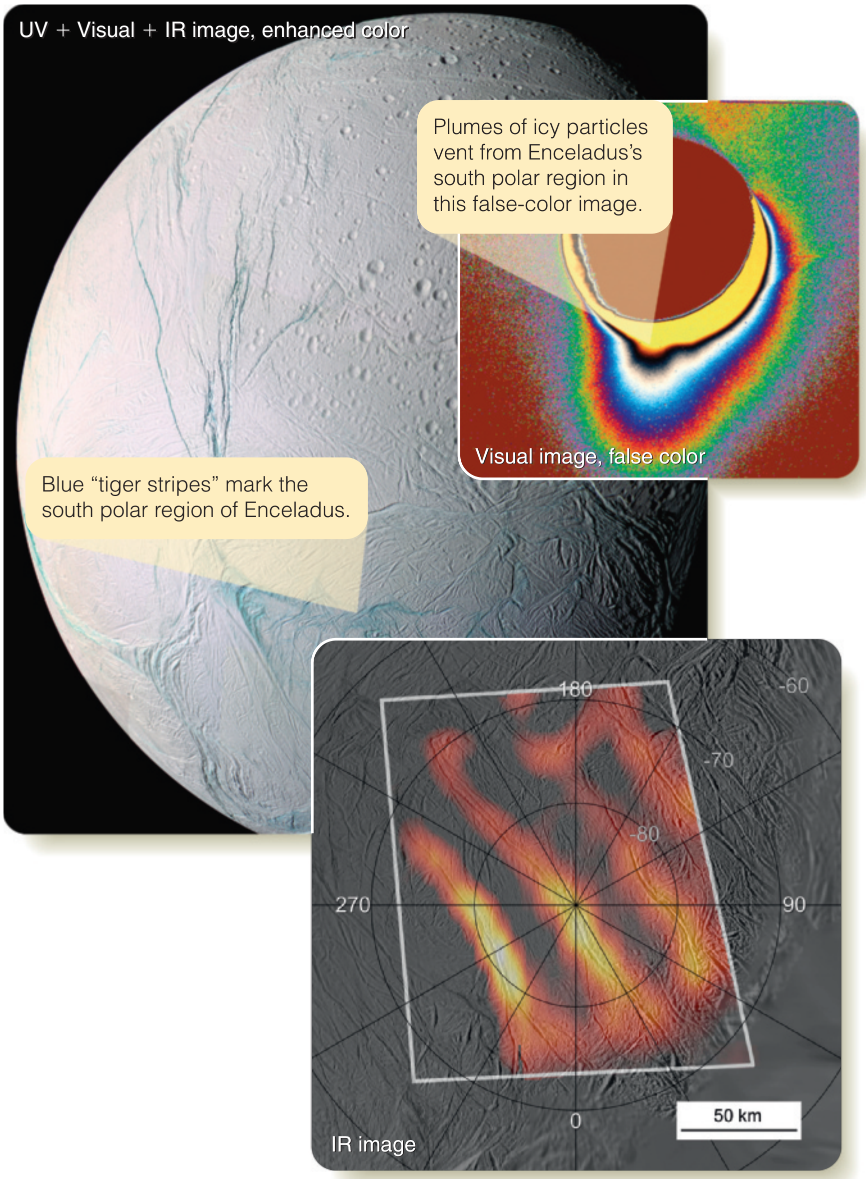


© Cengage Learning 2014

The *Cassini* spacecraft dropped the *Huygens* probe into the atmosphere of Titan. *Huygens* found, as expected, that Titan's surface is composed mainly of ices of water and methane at −180°C (−290°F). Images made by the descending probe showed dark drainage channels suggesting that liquid methane falls as rain, washes the dark goo off the higher terrain, and drains into the lowlands (Figure 11-9b). Such methane downpours may be rare, however. No direct evidence of liquid methane was detected as the probe descended, but later radar images made by the *Cassini* orbiter have detected what appear to be lakes presumably containing liquid methane (Figure 11-9d). Infrared images suggest the presence of methane volcanoes that replenish the methane in the atmosphere, so Titan must have some internal heat source to power the activity.

Most of the remaining moons of Saturn are small and icy, have no atmospheres, are heavily cratered and have dark, ancient surfaces. The moon Enceladus, however, shows signs of recent geological activity. Some parts of its surface contain 1000 times fewer craters than other regions, and infrared observations show that its south polar region is unusually warm and venting water and ice containing organic compounds (■ Figure 11-10). Evidently, a reservoir of liquid water lies just below the surface. At some point in its history, this moon may have been caught in a resonance with another moon and had its interior warmed by tidal heating.

Like nearly all moons in the solar system, Saturn's moons are tidally locked to their planet, rotating to keep the same side facing the planet. The leading side of these moons, the side facing forward in the orbit, is sometimes modified by debris. Iapetus, for example, has a cratered trailing side about as bright as dirty



© Cengage Learning 2014; Images: NASA/JPL/Space Science Institute

Figure 11-10

Saturn's moon Enceladus is venting water, ice, and organic molecules from geysers near its south pole. A thermal infrared image reveals internal heat leaking to space from the "tiger stripe" cracks where the geysers are located.

snow, but its leading side is as dark as fresh asphalt. One hypothesis is that the dark material is carbon-rich dust from meteoroid impacts on Phoebe, the next moon outward from Saturn. Iapetus also has a strange equatorial ridge that may have been produced by rapid rotation when the moon was young.

Saturn's Rings

The rings of Saturn are perhaps the most beautiful sight in our solar system. Study **The Ice Rings of Saturn** on pages 224–225 and notice three things:

1. The rings are made up of billions of ice particles, each in its own orbit around the planet. The ring particles you observe now can't be as old as Saturn. The rings must be replenished by impacts on Saturn's icy moons or other processes. The same is true of the rings around the other Jovian planets.

2. The gravitational effects of small moons called *shepherd satellites* can confine some rings in narrow strands or keep the edges of rings sharp. Moons can also produce waves in the rings that are visible as tightly wound ringlets.

3. The ring particles are confined in a thin plane spread among small moons and confined by gravitational interactions with larger moons. The rings of Saturn, and the rings of the other Jovian worlds, are created by and controlled by the planet's moons. Without the moons, there would be no rings.

Observations made by the *Cassini* spacecraft show that the ring particles have compositions that resemble that of Saturn's distant icy moon Phoebe. A large impact on Phoebe may be part of the complex history of Saturn's rings. Tiny icy particles from geysers on the moon Enceladus are understood to be the main source of material for the outermost low-density ring, discovered by the *Cassini* probe and named the E ring, within which Enceladus orbits.

The History of Saturn

Saturn formed in the outer solar nebula, where ice particles were stable and may have contained more trapped gases. The protoplanet grew rapidly and became massive enough to attract hydrogen and helium by gravitational collapse. The heavier elements sank to the middle to form a small core, and the hydrogen formed a liquid mantle containing liquid metallic hydrogen. The outward flow of heat from the interior drives convection inside the planet that helps produce its magnetic field. Because Saturn is smaller than Jupiter, the internal pressure is less, the planet contains less liquid metallic hydrogen, and its magnetic field is weaker.

The rings can't be *primordial.* That is, they can't be material left over from the formation of the planet. Such ices would have been vaporized and driven away by the heat of the protoplanet. Rather, you can suppose that the rings are debris from the occasional impacts of meteoroids, asteroids and comets on Saturn's icy moons.

Some of Saturn's moons are probably captured asteroids that wandered too close, but the larger moons almost certainly formed with Saturn. Many of Saturn's moons have ancient surfaces, but Enceladus and Titan have fresh surfaces and evidently ongoing geological activity. You will see a few more examples of recent geologic activity and also of captured moons when you explore farther from the sun.

SCIENTIFIC ARGUMENT

Why do the belts and zones on Saturn look so indistinct?

This argument compares Saturn with Jupiter. In a Jovian planet, the light-colored zones form where rising gas cools and condenses to form icy crystals of ammonia, which are visible as bright clouds. Saturn is twice as far from the sun as Jupiter, so sunlight is weaker and the atmosphere is colder. The gas in Saturn's atmosphere doesn't have to rise as high to reach temperatures cold enough to form clouds. Because the clouds form deeper in the hazy atmosphere, they are not as brightly illuminated by sunlight and look dimmer. Also, a layer of methane haze above the clouds makes the belts and zones look even less distinct.

You have used some simple physics to construct a logical argument that explains the hazy cloud features on Saturn. Now build a new argument. **Why do Saturn's rings have gaps and ringlets?**

11-4 Uranus

Now that you are familiar with the liquid giants in our solar system, you will be able to appreciate how strange the ice giants, Uranus and Neptune, are. Uranus, especially, seems to be the oddball of the family.

Uranus the Planet

Uranus is only one-third the diameter of Jupiter and only one-twentieth as massive. Four times farther from the sun than Jupiter, its atmosphere is almost 100°C colder than Jupiter's (**Celestial Profile 9**, page 230).

Uranus never grew massive enough to capture large amounts of gas from the nebula as Jupiter and Saturn did, so it has much less hydrogen and helium. Its internal pressure is enough less than Jupiter's that it should not contain any liquid metallic hydrogen. Models of Uranus based in part on its density and oblateness suggest that it has a small core of heavy elements and a deep mantle of partly solid water. Although that material is referred to as ice, it would not be anything like ice on Earth at the temperatures and pressures inside Uranus. The mantle also probably contains rocky material plus dissolved ammonia and methane. Circulation in that electrically conducting mantle may generate the planet's peculiar magnetic field, which is highly inclined to its axis of rotation. Above the mantle lies a deep hydrogen and helium atmosphere.

Uranus rotates on its side, with its equator inclined about 98° to its orbit. As a result the winter–summer contrast is extreme, with the sun passing near each of the planet's celestial poles at the solstices, so half of the planet is in perpetual darkness and the other half in perpetual light for the 21-year-long summer and winter seasons (■ Figure 11-11). Compare that with seasons on Earth, discussed in Chapter 3. When *Voyager 2* flew past in 1986, the planet's south pole was pointed almost directly at the sun.

Voyager 2 photos show a nearly featureless ball (■ Figure 11-12a). The atmosphere is mostly hydrogen and helium, but traces of methane absorb red light and thus make the atmosphere look green-blue. There is no belt–zone circulation visible in the *Voyager* photographs, although computer enhancements revealed a few clouds and bands around the south pole. In the decades since *Voyager 2* flew past Uranus, spring has come to the northern hemisphere of Uranus and autumn to the southern hemisphere. Images made by the Hubble Space Telescope and new large Earth-based telescopes reveal changing clouds and cloud bands in both hemispheres (Figure 11-12b).

Infrared measurements show that Uranus is radiating about the same amount of energy that it receives from the sun, meaning it has much less heat flowing out of its interior than Jupiter, Saturn, or Neptune. This may account for its limited atmospheric activity. Astronomers are not sure why Uranus differs in this respect from the other Jovian worlds.

Uranus's Moons

Until recently, astronomers could see only five moons orbiting Uranus. *Voyager 2* discovered ten more small moons in 1986, and more have been found in images recorded by powerful telescopes on Earth. The **International Astronomical Union (IAU)** that decides definitions and naming conventions for celestial objects and surface features has specified that moons of Uranus will continue to be named after characters in plays by William Shakespeare and Alexander Pope.

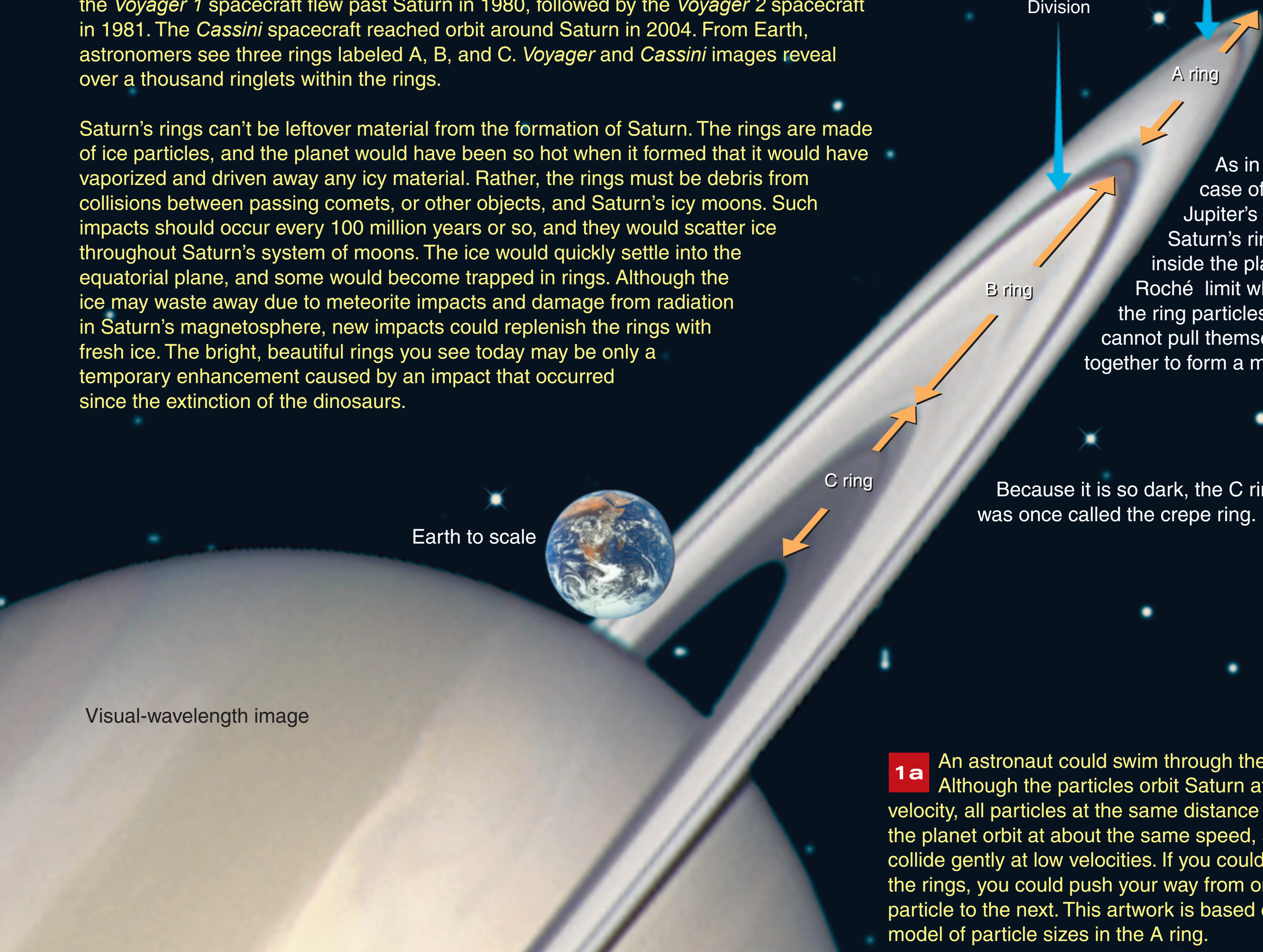
The five major moons of Uranus are all smaller than Earth's moon and have old, dark, cratered surfaces. A few have deep cracks, produced, perhaps, when the interior froze and expanded. In some cases, liquid water "lava" appears to have erupted and smoothed regions. Ariel is marked by broad, smooth-floored valleys that may have been cut by flowing ice (■ Figure 11-13a).

Miranda, the innermost moon, is only 14 percent the diameter of Earth's moon, but its surface is marked by grooves called **ovoids** (Figure 11-13b). These may have been caused by internal heat driving convection in the icy mantle. Rising currents of ice have deformed the crust and created the ovoids. By counting craters on the ovoids, astronomers conclude that the entire surface is old and the moon is no longer active. Perhaps it was warmed by tidal heating long ago.

The Ice Rings of Saturn

1 The brilliant rings of Saturn are made up of billions of ice particles ranging from microscopic specks to chunks bigger than a house. Each particle orbits Saturn in its own circular orbit. Much of what astronomers know about the rings was learned when the *Voyager 1* spacecraft flew past Saturn in 1980, followed by the *Voyager 2* spacecraft in 1981. The *Cassini* spacecraft reached orbit around Saturn in 2004. From Earth, astronomers see three rings labeled A, B, and C. *Voyager* and *Cassini* images reveal over a thousand ringlets within the rings.

Saturn's rings can't be leftover material from the formation of Saturn. The rings are made of ice particles, and the planet would have been so hot when it formed that it would have vaporized and driven away any icy material. Rather, the rings must be debris from collisions between passing comets, or other objects, and Saturn's icy moons. Such impacts should occur every 100 million years or so, and they would scatter ice throughout Saturn's system of moons. The ice would quickly settle into the equatorial plane, and some would become trapped in rings. Although the ice may waste away due to meteorite impacts and damage from radiation in Saturn's magnetosphere, new impacts could replenish the rings with fresh ice. The bright, beautiful rings you see today may be only a temporary enhancement caused by an impact that occurred since the extinction of the dinosaurs.



As in the case of Jupiter's ring, Saturn's rings lie inside the planet's Roché limit where the ring particles cannot pull themselves together to form a moon.

Because it is so dark, the C ring was once called the crepe ring.

Visual-wavelength image

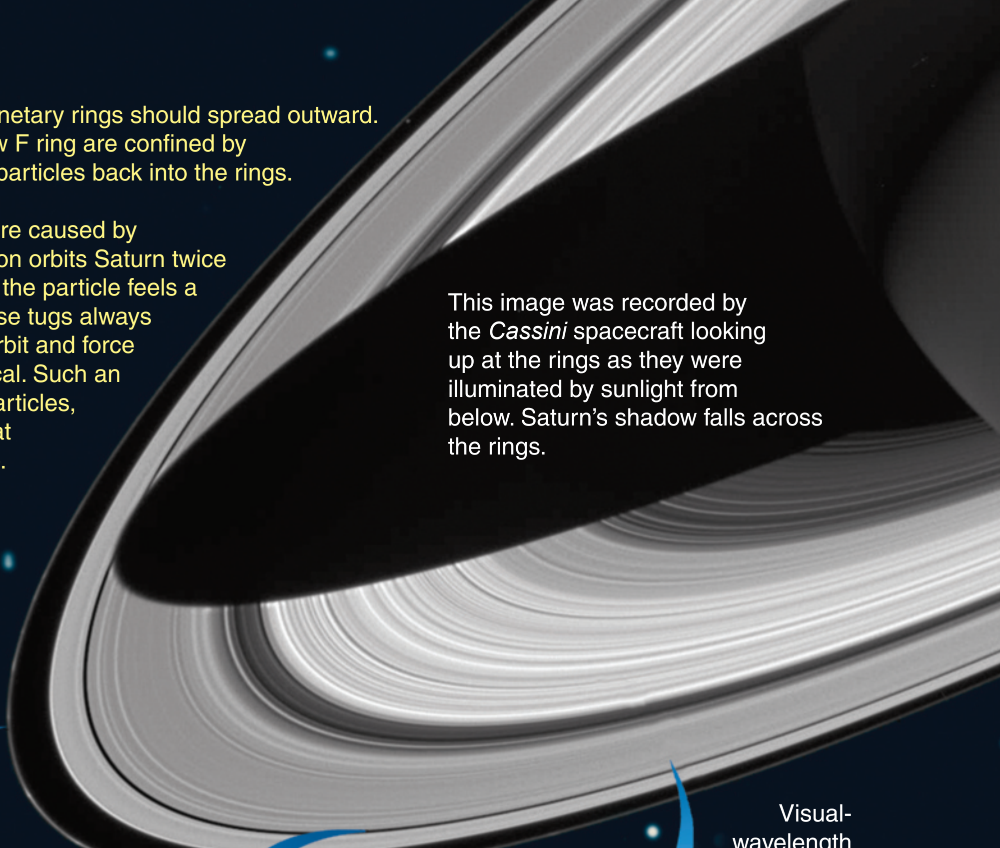
The C ring contains boulder-size chunks of ice, whereas most particles in the A and B rings are more like golf balls, down to dust-size ice crystals. Further, C ring particles are less than half as bright as particles in the A and B rings. *Cassini* observations show that the C ring particles contain less ice and more minerals.

1a An astronaut could swim through the rings. Although the particles orbit Saturn at high velocity, all particles at the same distance from the planet orbit at about the same speed, so they collide gently at low velocities. If you could visit the rings, you could push your way from one icy particle to the next. This artwork is based on a model of particle sizes in the A ring.

Line art on this page © Cengage Learning 2014; Images: NASA

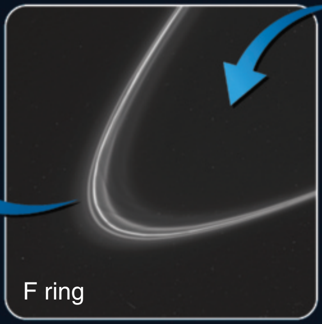
2 Because of collisions among ring particles, planetary rings should spread outward. The sharp outer edge of the A ring and the narrow F ring are confined by **shepherd satellites** that gravitationally usher straying particles back into the rings.

Some gaps in the rings, such as the Cassini Division, are caused by resonances with moons. A particle in the Cassini Division orbits Saturn twice for each orbit of the moon Mimas. On every other orbit, the particle feels a gravitational tug from Mimas. These tugs always occur at the same places in the orbit and force the orbit to become slightly elliptical. Such an orbit crosses the orbits of other particles, which results in collisions, and that removes the particle from the gap.

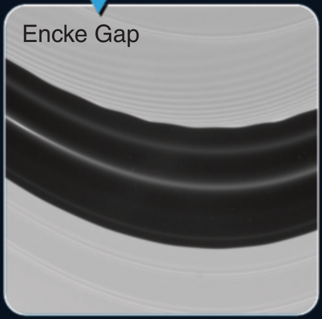


This image was recorded by the *Cassini* spacecraft looking up at the rings as they were illuminated by sunlight from below. Saturn's shadow falls across the rings.

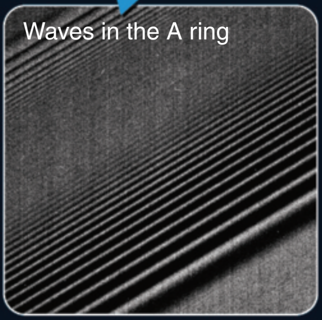
The F ring is clumpy and sometimes appears braided because of two shepherd satellites.



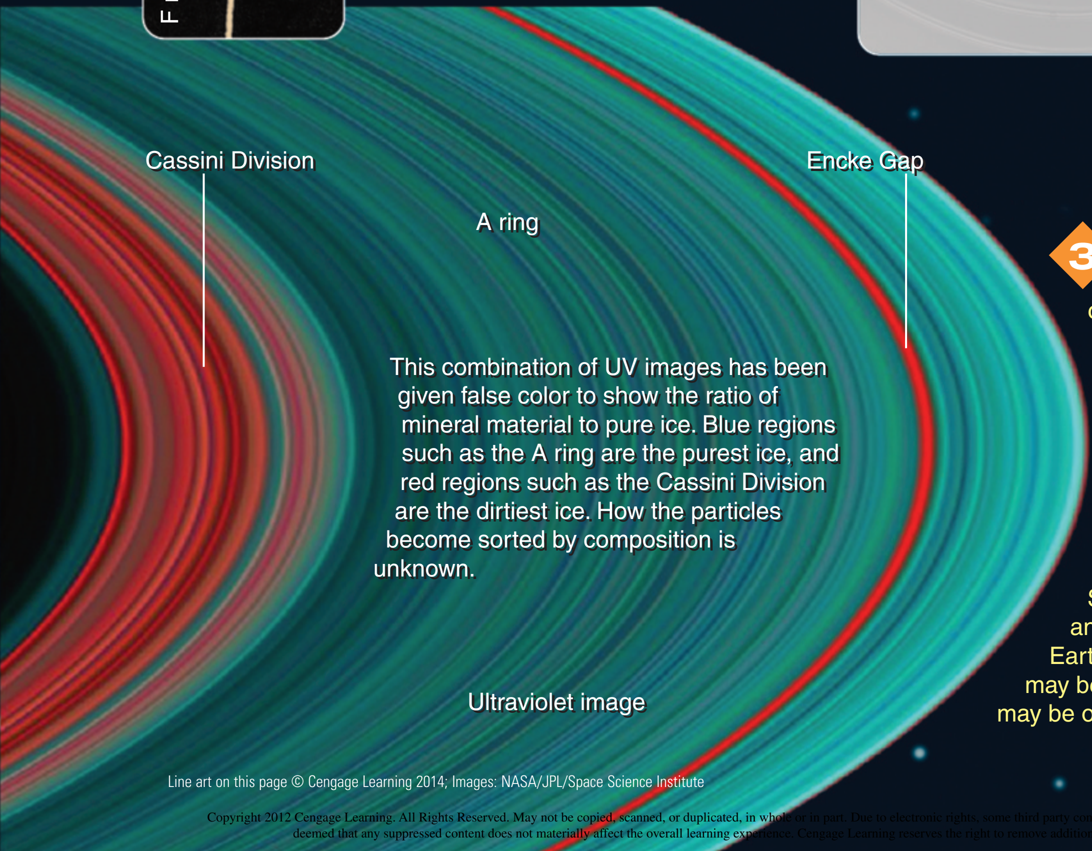
Visual-wavelength images



The Encke Gap is not empty. Note the ripples at the inner edge. A small moon orbits inside the gap.



Saturn does not have enough moons to produce all of its ringlets by resonances. Many are produced by tightly wound waves, much like the spiral arms found in disk galaxies.

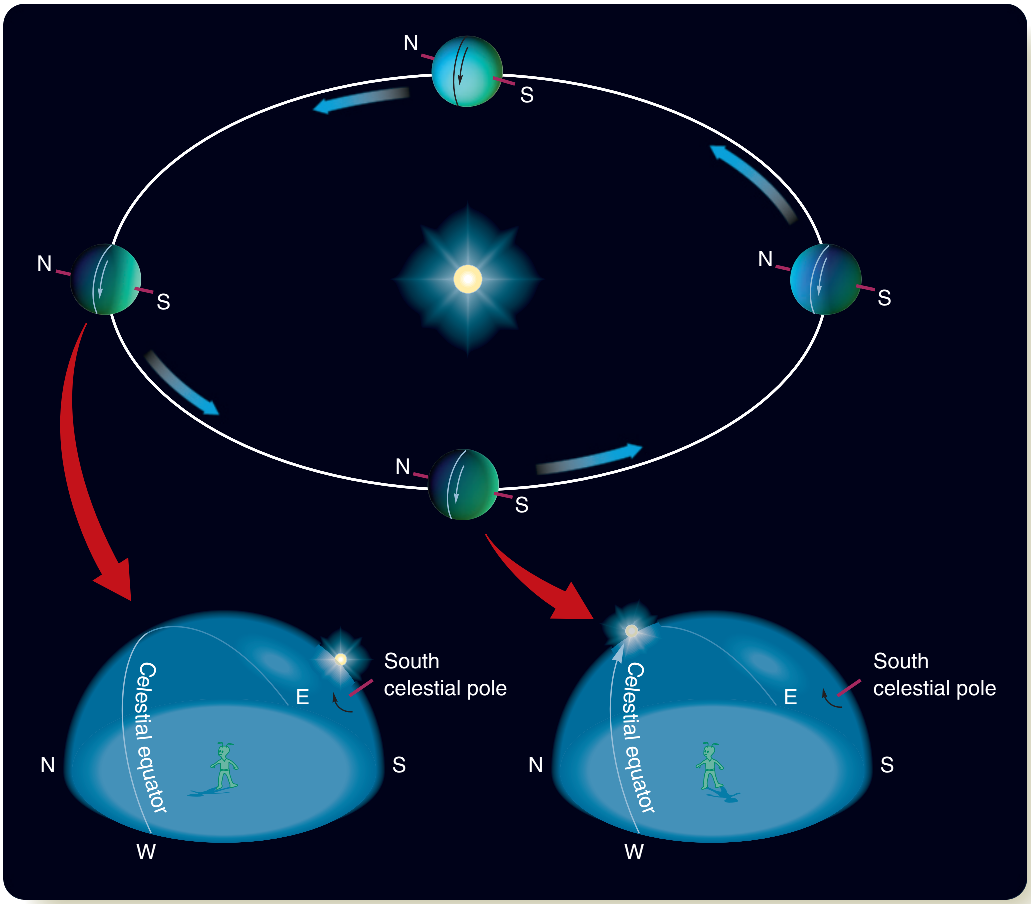


This combination of UV images has been given false color to show the ratio of mineral material to pure ice. Blue regions such as the A ring are the purest ice, and red regions such as the Cassini Division are the dirtiest ice. How the particles become sorted by composition is unknown.

3 How do moons happen to be at just the right places to confine the rings? That puts the cosmic cart before the horse. The ring particles get caught in the most stable orbits among Saturn's innermost moons. The rings push against the inner moons, but those moons are locked in place by resonances with larger, outer moons. Without the moons, the rings would spread and dissipate.

Saturn's rings are a very thin layer of particles and nearly vanish when the rings turn edge-on to Earth. Although ripples in the rings caused by waves may be hundreds of meters high, the sheet of particles may be only about ten meters thick.

Line art on this page © Cengage Learning 2014; Images: NASA/JPL/Space Science Institute



© Cengage Learning 2014

Figure 11-11

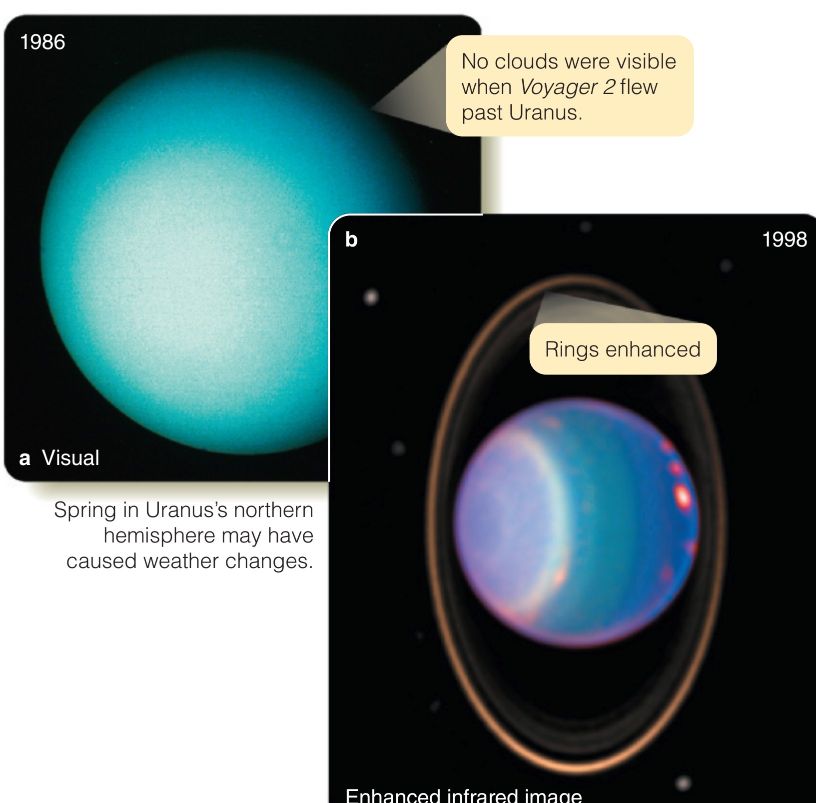
Uranus rotates on an axis that is tipped about 98° from the perpendicular to its orbit, so its seasons are extreme. When one of its poles is pointed nearly at the sun (a solstice), an inhabitant of Uranus would see the sun near a celestial pole, never rising or setting. As Uranus orbits the sun, the planet maintains the direction of its axis in space, and thus the sun moves from pole to pole. At the time of an equinox on Uranus, the sun would be on the celestial equator and would rise and set with each rotation of the planet. Compare with similar diagrams for Earth on page 27.

Uranus's Rings

The rings of Uranus are not easily visible from Earth. The first hint that Uranus has rings came from **occultations,** the passage of the planet in front of a star during which the rings momentarily blocked the star's light, observed by astronomers onboard the Kuiper Airborne Observatory in 1977. Most of what astronomers know about these rings comes from the observations by the *Voyager 2* spacecraft. Their composition appears to be water ice mixed with methane that has been darkened by exposure to solar wind particles trapped in the planet's magnetosphere.

Study **The Rings of Uranus and Neptune** on pages 228–229 and notice three important points:

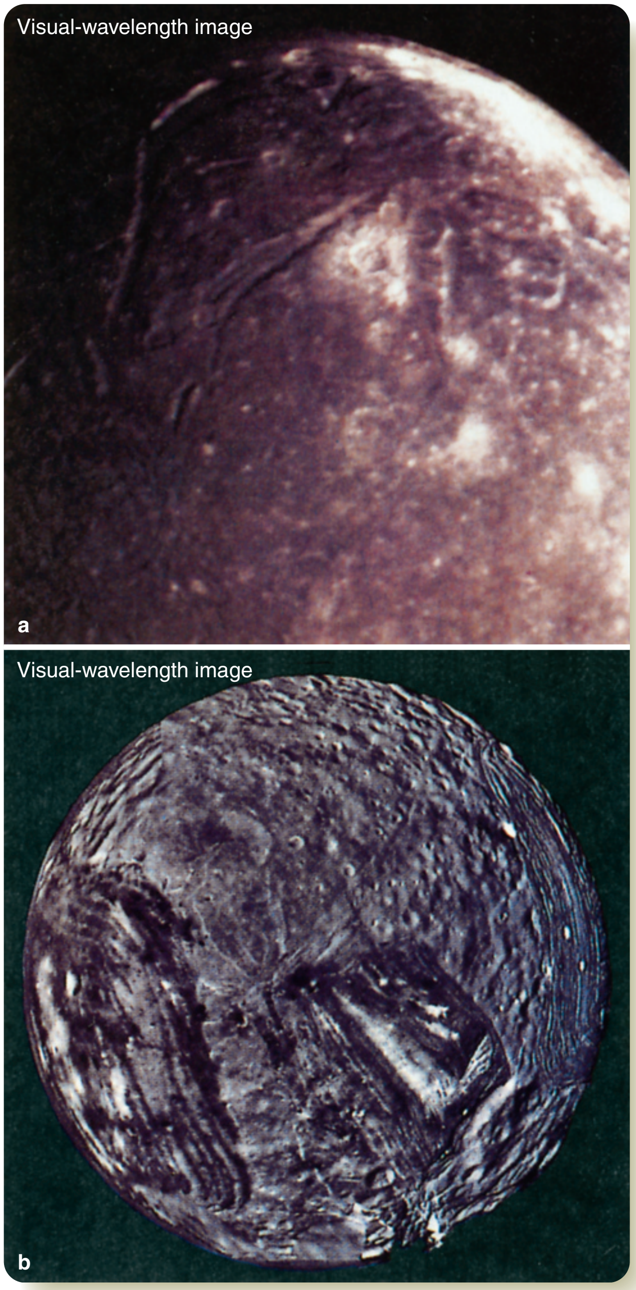
1. The rings of Uranus were discovered during an *occultation* when Uranus crossed in front of a star.
2. The rings are dark, contain little dust, and are confined by small moons.
3. Particles orbiting in the rings around Uranus and Neptune cannot survive for long periods, so the rings need to be resupplied with material from impacts on moons, as is also true for the rings around Jupiter and Saturn.



© Cengage Learning 2014; Images: NASA and Erich Karkoschka, University of Arizona

Figure 11-12

(a) If you had been riding onboard *Voyager* as it passed Uranus, the planet would have looked like a bland green-blue ball. (b) Later ground-based contrast-enhanced images reveal traces of belt–zone circulation deep in the atmosphere.



NASA

NASA

Figure 11-13

Evidence of geological activity on two Uranian moons. (a) Ariel has an old cratered surface, but some regions are marked by broad shallow valleys with few craters. (b) The face of Miranda is marred by ovoids, which are believed to have formed when internal heating caused slow convection in the ice of the moon's mantle. Note the 5-km-high cliff at the lower edge of the moon.

When you read about Neptune's rings later in this chapter, you can return to this art spread and see how closely the two ring systems compare.

In 2006, astronomers found two new, very faint rings orbiting far outside the previously known rings of Uranus. The newly discovered satellite Mab appears to be the source of particles for the larger ring, and the smaller of the new rings is confined between the orbits of the moons Portia and Rosalind.

A History of Uranus

Uranus never grew massive enough to capture large amounts of gas from the solar nebula as did Jupiter and Saturn. Uranus is rich in water and ice rather than in hydrogen and helium.

Modern models of the origin of the solar system suggest that Uranus and Neptune formed closer to the sun than their present positions. Interactions with massive Jupiter and Saturn could have gradually moved Uranus and Neptune outward, and tidal effects may have produced Uranus's peculiar rotation. Another hypothesis is that Uranus was struck by a large planetesimal as it was forming and given its highly inclined rotation.

The highly inclined magnetic field of Uranus may be produced by convection in its electrically conducting mantle. With very little heat flowing out of the interior, this convection must be limited.

SCIENTIFIC ARGUMENT

Why are the rings of Uranus so narrow?

Unlike the rings of Jupiter and Saturn, the rings of Uranus are quite narrow, like hoops of wire. You would expect collisions among ring particles to gradually spread the rings out into thin sheets, so something must be confining the narrow rings. In fact, two small moons have been found orbiting just inside and outside the ε ring. If a ring particle drifts away from the ring, the corresponding moon's gravity will nudge it back into the ring, so they are called shepherd satellites. More shepherd satellites, too small to have been detected so far, are thought to control the other rings. Thus, the rings of Uranus resemble Saturn's narrow F ring.

Now expand your argument. **How do moons happen to be in the right place to keep the rings narrow?**

11-5 Neptune

Through a telescope on Earth, Neptune looks like a tiny blue dot with no visible cloud features. In 1989, *Voyager 2* flew past and revealed most of what we know today about Neptune.

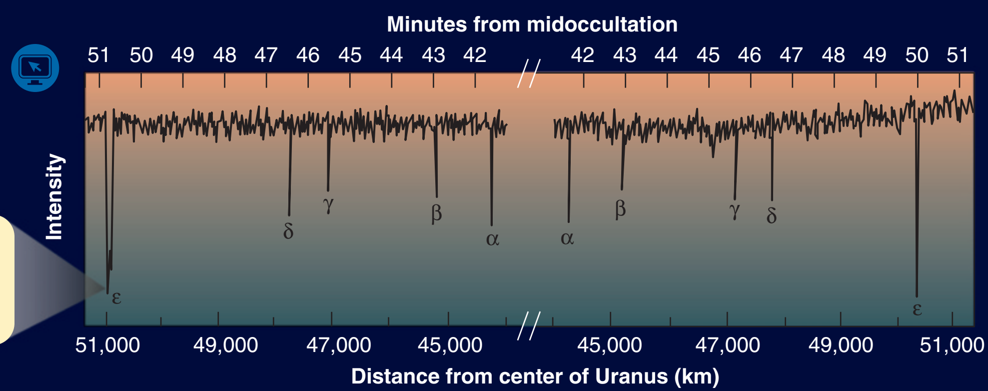
Neptune the Planet

Almost exactly the same size as Uranus, Neptune has a similar interior. Model calculations predict that a small core of heavy elements lies within a slushy mantle of water, ices, and rocky materials below a hydrogen-rich atmosphere (**Celestial Profile 10**, page 230). Yet Neptune looks quite different on the outside from Uranus; Neptune is dramatically blue and has active cloud formations. Neptune's dark blue tint is caused by its atmospheric composition of one and a half times more methane than Uranus. Methane absorbs red photons better than blue and scatters blue

The Rings of Uranus and Neptune

1 The rings of Uranus were discovered in 1977, when Uranus crossed in front of a star. During this **occultation**, astronomers saw the star dim a number of times before and again after the planet crossed over the star. The dips in brightness were caused by rings circling Uranus.

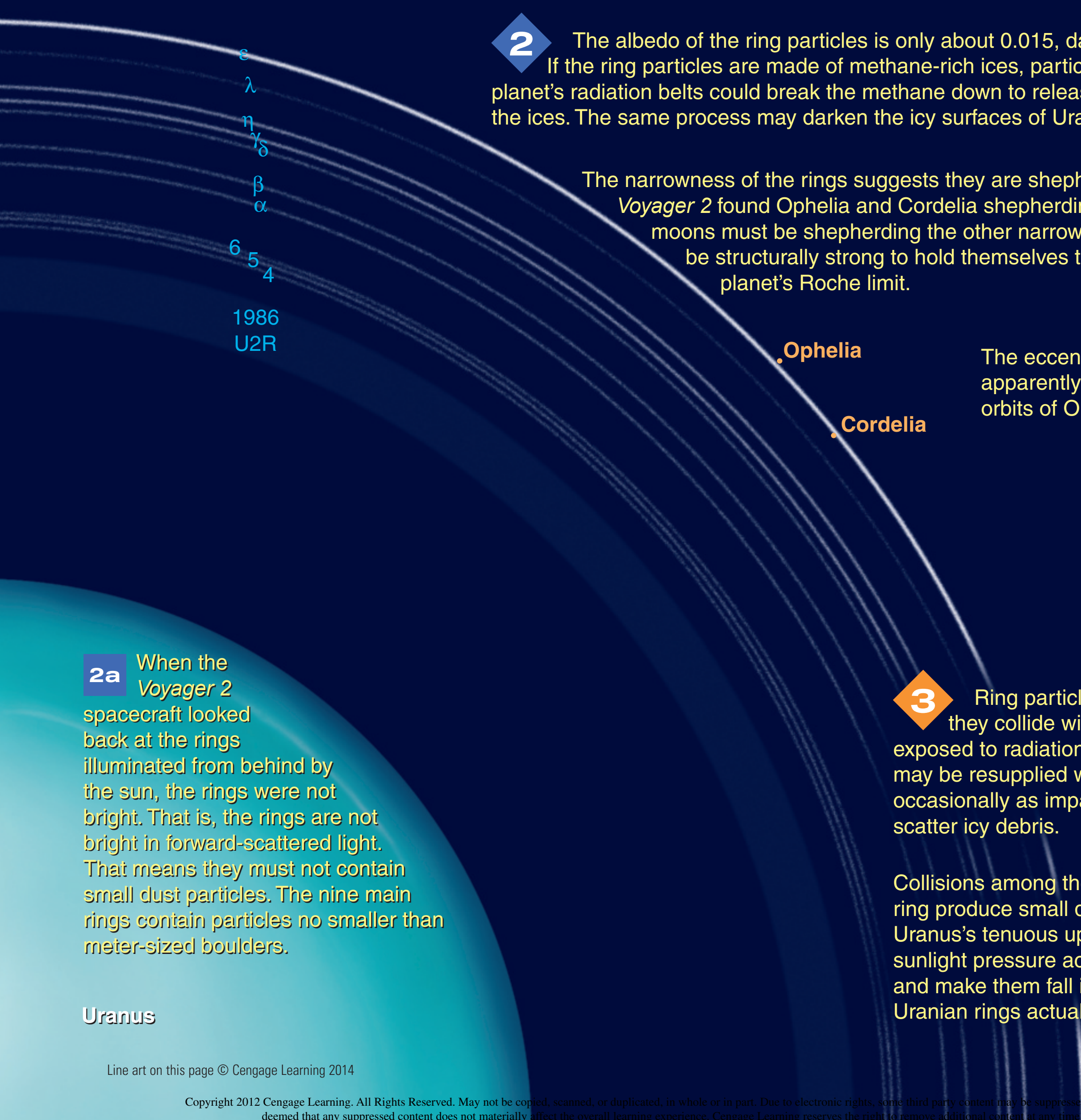
More rings were discovered by *Voyager 2*. The rings are identified in different ways depending on when and how they were discovered.



2 The albedo of the ring particles is only about 0.015, darker than lumps of coal. If the ring particles are made of methane-rich ices, particle radiation from the planet's radiation belts could break the methane down to release carbon and darken the ices. The same process may darken the icy surfaces of Uranian moons.

The narrowness of the rings suggests they are shepherded by small moons. *Voyager 2* found Ophelia and Cordelia shepherding the ε ring. Other small moons must be shepherding the other narrow rings. Such moons must be structurally strong to hold themselves together inside the planet's Roche limit.

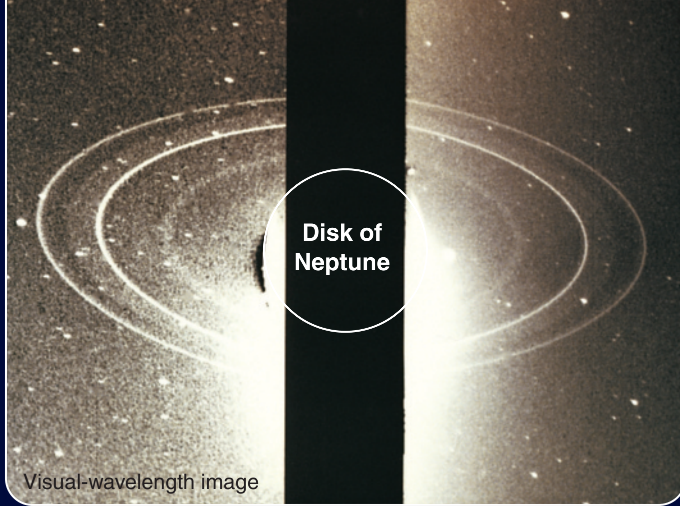
The eccentricity of the ε ring is apparently caused by the eccentric orbits of Ophelia and Cordelia.



2a When the *Voyager 2* spacecraft looked back at the rings illuminated from behind by the sun, the rings were not bright. That is, the rings are not bright in forward-scattered light. That means they must not contain small dust particles. The nine main rings contain particles no smaller than meter-sized boulders.

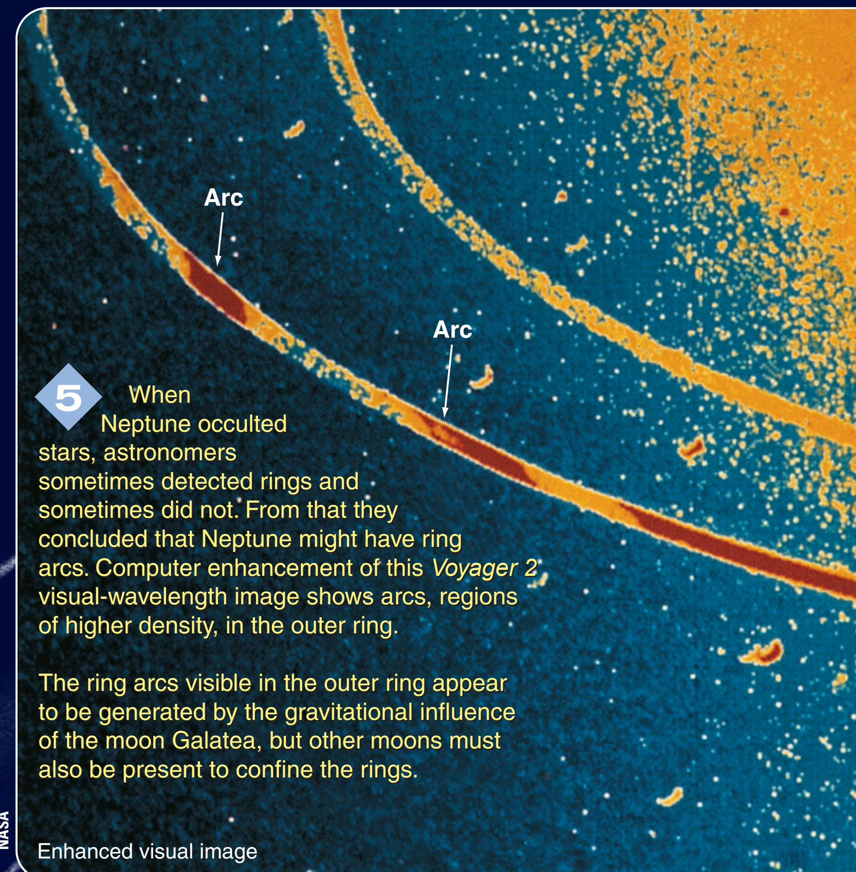
3 Ring particles don't last forever as they collide with each other and are exposed to radiation. The rings of Uranus may be resupplied with fresh particles occasionally as impacts on icy moons scatter icy debris.

Collisions among the large particles in the ring produce small dust grains. Friction with Uranus's tenuous upper atmosphere plus sunlight pressure act to slow the dust grains and make them fall into the planet. The Uranian rings actually contain very little dust.



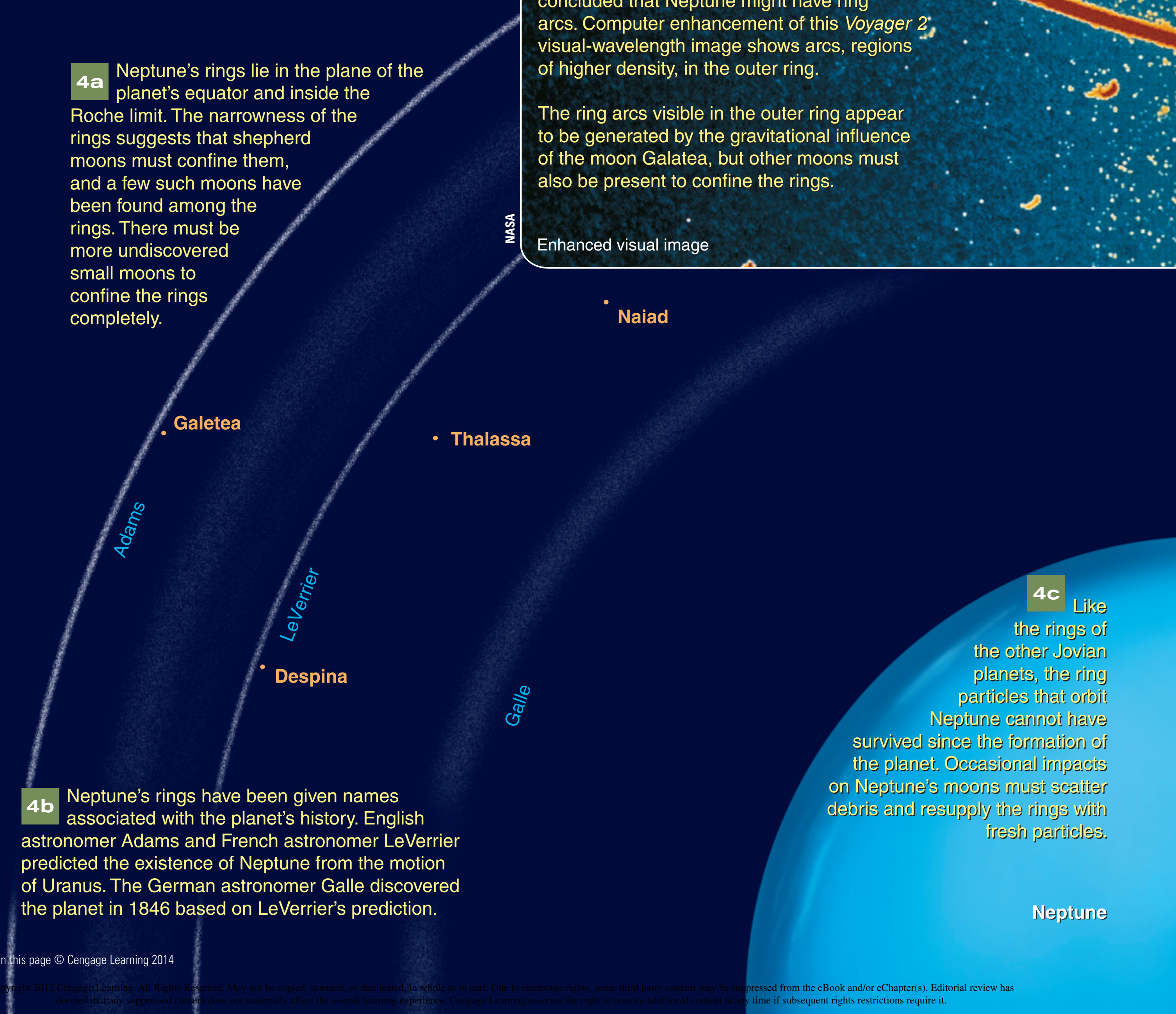
The rings of Neptune are bright in forward-scattered light, as in the image above, and that indicates that the rings contain significant amounts of dust. The ring particles are as dark as those that circle Uranus, so they probably also contain methane-rich ice darkened by radiation.

4 The brightness of Neptune is hidden behind the black bar in this *Voyager 2* image. Two narrow rings are visible, and a wider, fainter ring lies closer to the planet. More ring material is visible between the two narrow rings.



5 When Neptune occulted stars, astronomers sometimes detected rings and sometimes did not. From that they concluded that Neptune might have ring arcs. Computer enhancement of this *Voyager 2* visual-wavelength image shows arcs, regions of higher density, in the outer ring.

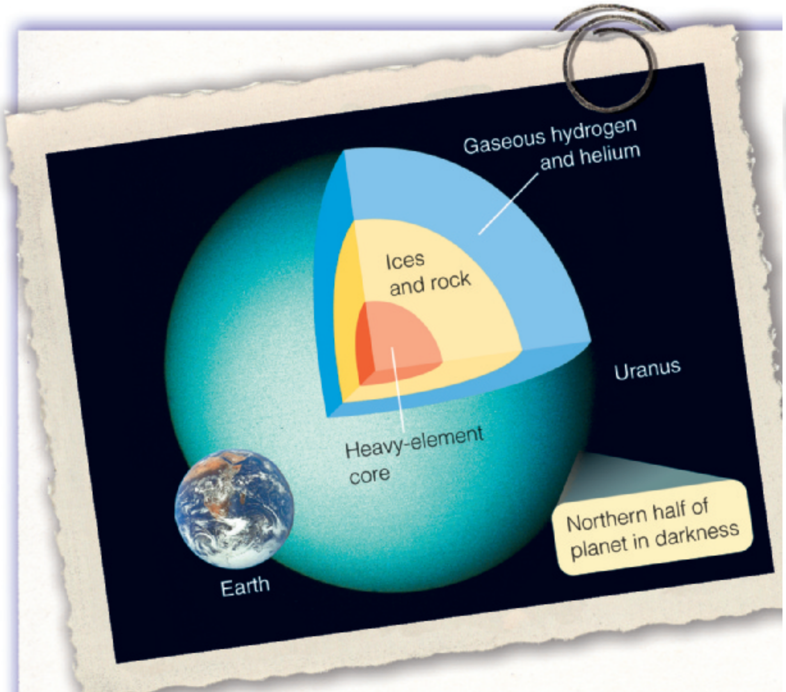
The ring arcs visible in the outer ring appear to be generated by the gravitational influence of the moon Galatea, but other moons must also be present to confine the rings.



4a Neptune's rings lie in the plane of the planet's equator and inside the Roche limit. The narrowness of the rings suggests that shepherd moons must confine them, and a few such moons have been found among the rings. There must be more undiscovered small moons to confine the rings completely.

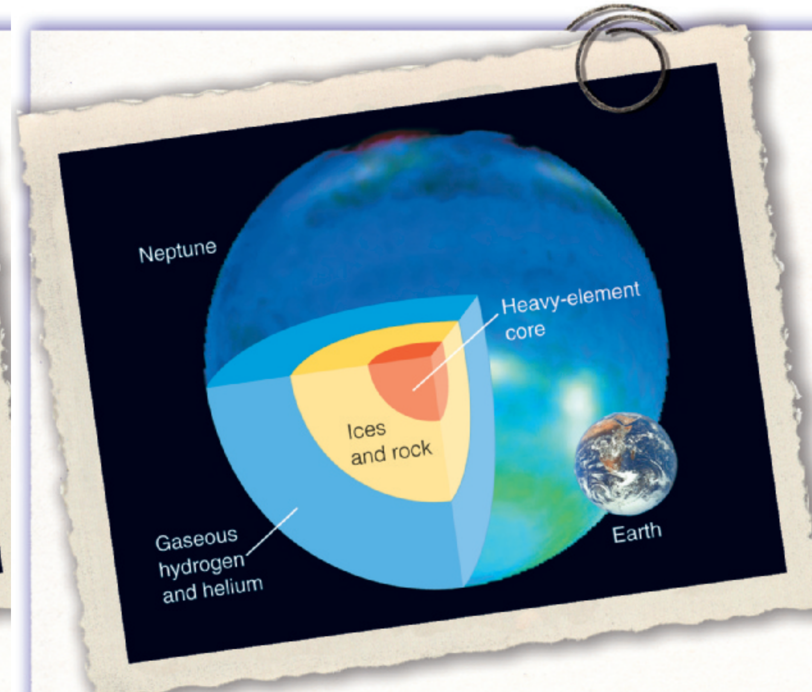
4b Neptune's rings have been given names associated with the planet's history. English astronomer Adams and French astronomer LeVerrier predicted the existence of Neptune from the motion of Uranus. The German astronomer Galle discovered the planet in 1846 based on LeVerrier's prediction.

4c Like the rings of the other Jovian planets, the ring particles that orbit Neptune cannot have survived since the formation of the planet. Occasional impacts on Neptune's moons must scatter debris and resupply the rings with fresh particles.



© Cengage Learning 2014; Image: NASA

Uranus rotates on its side, and, when Voyager 2 *flew past in 1986, the planet's south pole was pointed almost directly at the sun.*



© Cengage Learning 2014; Image: NASA

Neptune was tipped slightly away from the sun when the Hubble Space Telescope recorded this image. The interior is much like that of Uranus, but Neptune has more heat flowing outward.

Celestial Profile 9: Uranus

Motion:

Average distance from the sun	19.2 AU (2.88×10^9 km)
Eccentricity of orbit	0.044
Inclination of orbit to ecliptic	0.8°
Orbital period	84.3 y
Period of rotation	17.23 h
Inclination of equator to orbit	97.8° (retrograde rotation)

Characteristics:

Equatorial diameter	5.11×10^4 km (4.01 $D_{\oplus}$)
Mass	8.68×10^{25} kg (14.5 $M_{\oplus}$)
Average density	1.27 g/cm^3
Gravity	0.9 Earth gravity
Escape velocity	21 km/s (1.9 $V_{\oplus}$)
Temperature at cloud tops	55°K (−360°F)
Albedo	0.30
Oblateness	0.023

Personality Point:

Uranus was discovered in 1781 by William Herschel, a German-born scientist who lived and worked most of his life in England. He named the new planet *Georgium Sidus*, meaning "George's Star" in Latin, after the English King George III. European astronomers, especially the French, refused to accept a planet named after an English king. Instead, they called the planet Herschel. Years later, German astronomer J. E. Bode suggested it be named Uranus after the oldest of the Greek gods.

Celestial Profile 10: Neptune

Motion:

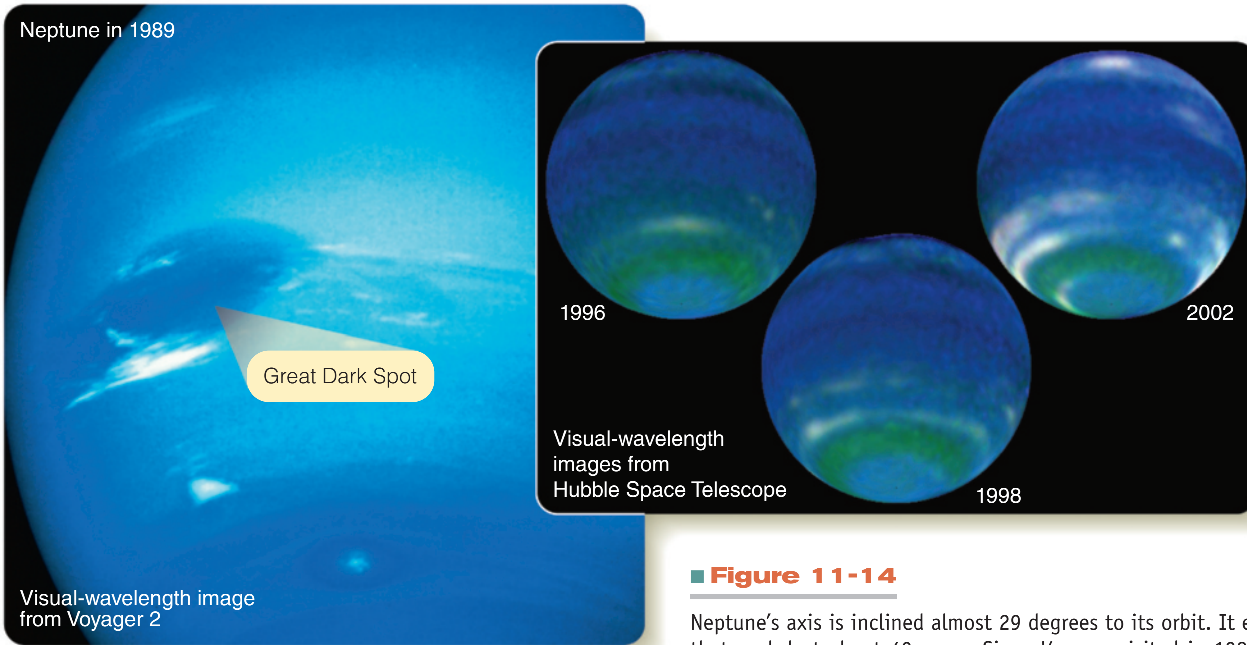
Average distance from the sun	30.1 AU (4.50×10^9 km)
Eccentricity of orbit	0.011
Inclination of orbit to ecliptic	1.8°
Orbital period	164.8 y
Period of rotation	16.11 h
Inclination of equator to orbit	28.3°

Characteristics:

Equatorial diameter	4.95×10^4 km (3.88 $D_{\oplus}$)
Mass	1.02×10^{26} kg (17.1 $M_{\oplus}$)
Average density	1.64 g/cm^3
Gravity	1.1 Earth gravities
Escape velocity	24 km/s (2.1 $V_{\oplus}$)
Temperature at cloud tops	55°K (−360°F)
Albedo	0.29
Oblateness	0.017

Personality Point:

A British and a French astronomer independently calculated the existence and location of Neptune from its gravitational influence on the motion of Uranus. British observers were too slow to act on this information; Neptune was discovered in 1846, and the French astronomer got the credit. Because of its blue color, astronomers named Neptune after the god of the sea.



© Cengage Learning 2014; Images: NASA, L. Sromovsky, and P. Fry, University of Wisconsin-Madison

■ Figure 11-14

Neptune's axis is inclined almost 29 degrees to its orbit. It experiences seasons that each last about 40 years. Since *Voyager* visited in 1989, spring has come to the southern hemisphere, and the weather has clearly changed, which is surprising because sunlight at Neptune is 900 times dimmer than at Earth.

photons better than red, giving Neptune a blue color and Uranus a green-blue color.

Atmospheric circulation on Neptune is much more dramatic than on Uranus. When *Voyager 2* flew by Neptune in 1989, the largest feature was the Great Dark Spot (■ Figure 11-14). Roughly the size of Earth, the spot seemed to be an atmospheric circulation pattern much like Jupiter's Great Red Spot. Smaller spots were visible in Neptune's atmosphere, and photos showed they were circulating like hurricanes. More recently, the Hubble Space Telescope has photographed Neptune and found that the Great Dark Spot is gone and new cloud formations have appeared, as shown in Figure 11-14. Evidently, weather patterns on Neptune are very changeable.

The atmospheric activity on Neptune is apparently driven by heat flowing from the interior plus some contribution by dim sunlight 30 AU from the sun. The heat causes convection in the atmosphere, which the rapid rotation of the planet converts into high-speed winds, high-level white clouds of methane ice crystals, and rotating storms visible as spots. Neptune may have more activity than Uranus because it has more heat flowing out of its interior, for reasons that are unclear.

Like Uranus, Neptune has a highly inclined magnetic field that must be linked to circulation in the interior. In both cases, astronomers suspect that ammonia dissolved in the liquid water mantle makes the mantle a good electrical conductor and that convection in the water, coupled with the rotation of the planet, drives the dynamo effect and generates the magnetic field.

Neptune's Moons

Neptune has two moons that were discovered from Earth before *Voyager 2* flew past in 1989. *Voyager* discovered six more very small moons. Since then, a few more small moons have been found by astronomers using Earth-based telescopes.

The two largest moons have peculiar orbits. Nereid, about a tenth the size of Earth's moon, follows a large, elliptical orbit, taking nearly an Earth year to circle Neptune once. Triton, almost 80 percent the size of Earth's moon, orbits Neptune retrograde (backward). These odd orbits suggest that the system was disturbed long ago in an interaction with some other body, such as a massive planetesimal.

With a surface temperature of about 36 K (−395°F), Triton has an atmosphere of nitrogen and methane about 10^5 times less dense than Earth's. A significant part of Triton is ice, and deposits of nitrogen frost are visible at the southern pole (■ Figure 11-15a), which, at the time *Voyager 2* flew past, had been turned toward sunlight for 30 years. The nitrogen frost appears to be vaporizing in the sunlight and is probably refreezing in the darkness at Triton's north pole.

Many features on Triton suggest it has had an active past. It has few craters on its surface, but it does have long faults that appear to have formed when the icy crust broke. Some approximately round basins about 400 km in diameter appear to have been flooded time after time by liquids from the interior (Figure 11-15b). Analysis of dark smudges visible in the southern polar cap (Figure 11-15a) reveals that these are deposits

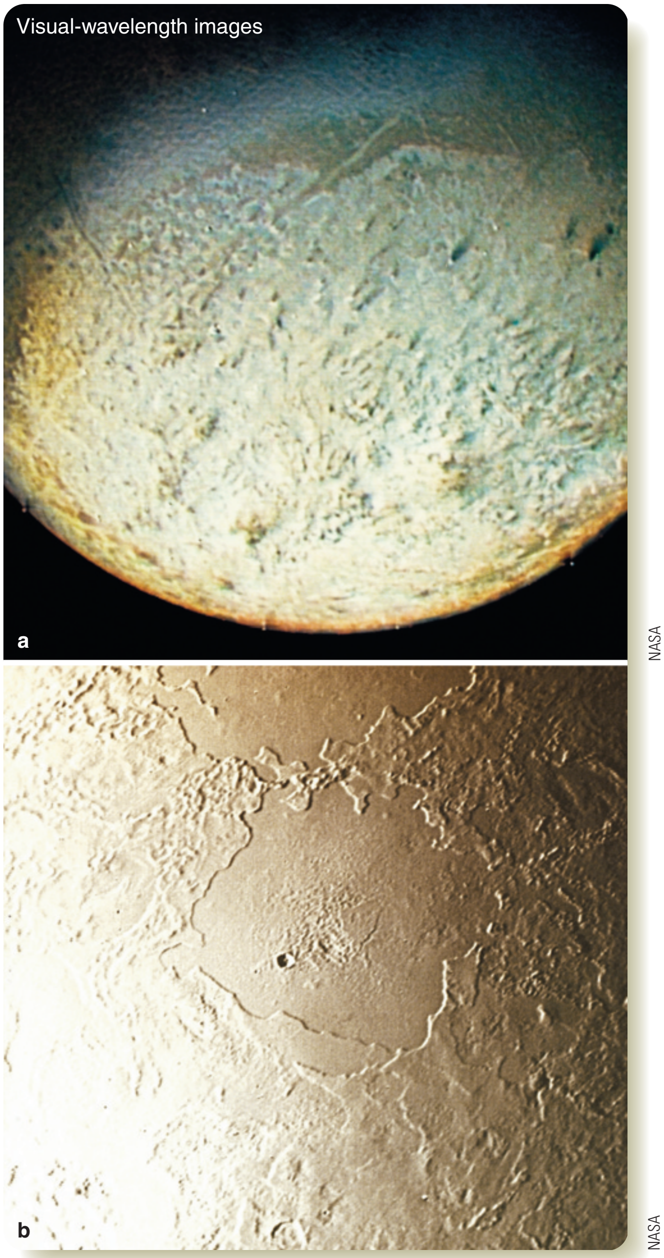


Figure 11-15

Visible-wavelength images of Neptune's moon Triton. (a) Triton's southern polar cap is formed of nitrogen frost. Note dark smudges caused by organic compounds sprayed from nitrogen geysers, and the absence of craters. (b) These round basins on Triton appear to have been repeatedly flooded by liquid from the interior.

produced when liquid nitrogen in the crust, warmed by the sun, erupts through vents and spews up to 8 km (5 mi) high into the atmosphere. Methane in the gas is converted by solar ultraviolet radiation into dark compounds that fall back, leaving black smudges.

By counting craters on Triton, planetary scientists conclude that the surface has been active as recently as a million years ago and may still be active. The energy source for Triton's geyser activity could come from radioactive decay. The moon is two-thirds rock, and although such a small world would not be able to generate sufficient radioactive decay to keep molten rock flowing to its surface, frigid Triton may be the site of water–ammonia volcanism: A mixture of water and ammonia could melt at very low temperatures and erupt to resurface parts of the moon.

Neptune's Rings

Neptune's rings are faint and very hard to detect from Earth, but they illustrate some interesting principles of comparative planetology.

Look again at **The Rings of Uranus and Neptune** on pages 228–229 and compare the rings of Neptune with those of Uranus. Notice two additional points:

4 Neptune's rings, named after the astronomers involved in the discovery of the planet, are similar to those of Uranus but contain more small dust particles.

5 Also, Neptune's rings show another way that moons can interact with rings: One of Neptune's moons is producing short arcs in the outermost ring.

Neptune's rings resemble the rings of Uranus, Saturn, and Jupiter in one important way. As you have already learned, these rings can't be primordial. That is, they can't have lasted since the formation of the planets. Planetary rings are constantly being remade.

The History of Neptune

Neptune must have formed much as Uranus did, growing slowly and never becoming massive enough to trap large amounts of hydrogen and helium. It developed a core of heavy elements, a mantle of slushy ices and rock, and a deep hydrogen-rich atmosphere. Neptune's internal heat may be generated partly by radioactive decay in its core and partly by dense material sinking inward. Laboratory experiments show that at the temperatures and pressures expected deep in the atmospheres of Neptune and Uranus, methane can decompose, and the released carbon might form diamond crystals perhaps as large as pebbles. A continuous flow of diamonds falling into a planet's interior would release gravitational energy and help warm the planet. This process may be the source of some of Neptune's internal heat, but the lack of internal heat in Uranus remains a puzzle. (The possibility of a planetwide hailstorm of diamonds serves to warn you that other worlds are truly un-Earthly and may harbor things you can hardly imagine.) The heat flowing outward toward Neptune's surface can drive convection, produce a magnetic field, and help create atmospheric circulation.

The moons of Neptune suggest some cataclysmic encounter long ago that put Nereid into a long-period elliptical orbit and Triton into a retrograde orbit. You have seen evidence of major impacts throughout the solar system, so such interactions may have been fairly common. Certainly, impacts on the satellites could provide the debris that is trapped among the smaller moons to form the rings.

SCIENTIFIC ARGUMENT

Why is Neptune blue?

In building this argument, you need to be careful not to be misled by the words you use. When you look at something, you really turn your eyes toward it and receive light from the object. The light Earth receives from Neptune is sunlight that was scattered from various layers of Neptune and journeys back to your eyes. Sunlight entering Neptune's atmosphere must pass through hydrogen gas that contains a small amount of methane, which is a good absorber of longer wavelengths. As a result, red photons are more likely to be absorbed than blue photons, and that makes the light bluer. Furthermore, when the light is scattered in deeper layers, the shorter-wavelength photons are most likely to be scattered, and thus the light that finally emerges from the atmosphere and reaches your telescope is poor in longer wavelengths. It looks blue.

This argument shows how a careful, step-by-step analysis of a natural process can help you better understand how nature works. Now expand your argument. **Why do the clouds on Neptune look white?**

11-6 Pluto and the Kuiper Belt

Out on the edge of the solar system orbits a family of small, icy worlds. Pluto was the first to be discovered, in 1930, but modern telescopes have found more.

You may have learned in school that there are nine planets in our solar system, but in 2006 the IAU voted to remove Pluto from the list of planets and reclassify it as a "dwarf planet." Pluto is a very small, icy world: It isn't Jovian, and it isn't Terrestrial. Its orbit is highly inclined and elliptical enough that Pluto actually comes closer to the sun than Neptune at times. To understand Pluto's status, you can use comparative planetology to analyze Pluto and then compare it with its neighbors.

Pluto is very difficult to observe from Earth. It has only about two-thirds the diameter of Earth's moon. In Earth-based telescopes, it never looks like more than a faint point of light, and even in *Hubble Space Telescope* images it shows little detail. Orbiting so far from the sun, Pluto is cold enough to freeze most compounds you think of as gases, and spectroscopic observations have found evidence of nitrogen ice on its surface. Pluto has a thin atmosphere of nitrogen and carbon monoxide with small amounts of methane.

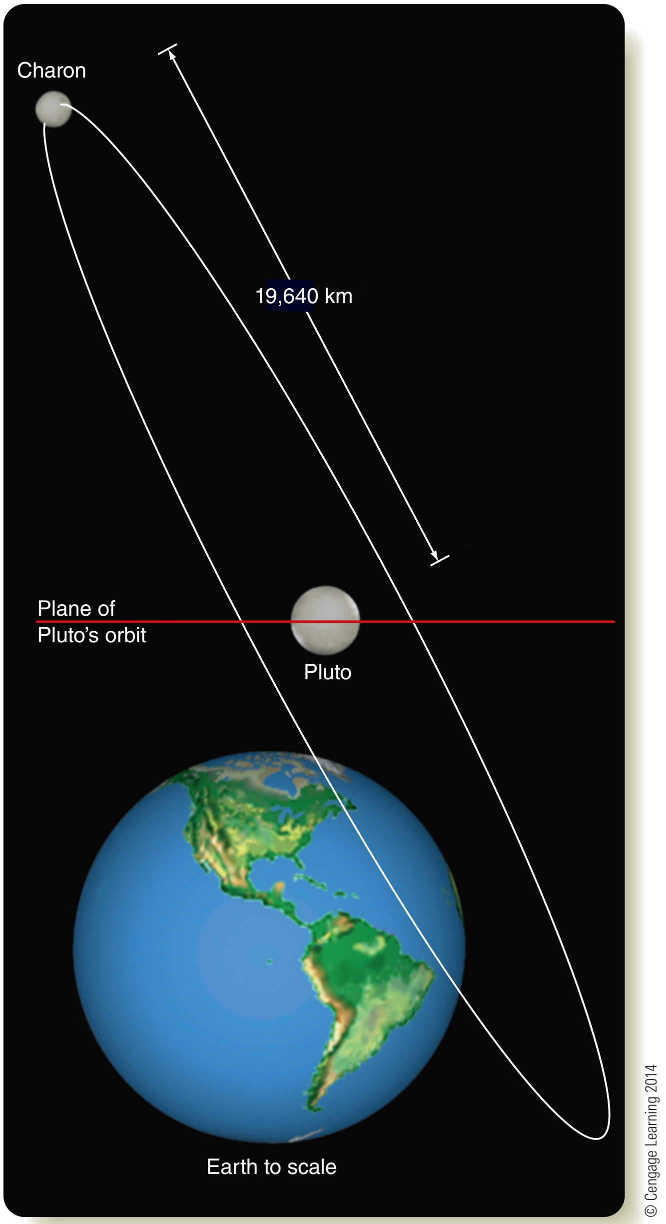


Figure 11-16

Pluto, Pluto's moon Charon, and Charon's orbit are pictured here in scale relative to Earth. Charon's orbit is tipped 118° to the plane of Pluto's orbit, making its motion retrograde.

At this writing, Pluto has five known moons. Four of them are quite small, but Charon is relatively large, with half of Pluto's diameter. Charon orbits Pluto with a period of 6.4 days in an orbit highly inclined to the ecliptic (Figure 11-16). Pluto and Charon are tidally locked to face each other, so Pluto's axis of rotation is also highly inclined.

Charon's orbit size and period plus Kepler's third law reveal that the mass of the system is only about 0.002 Earth mass. Most of that mass is Pluto, which has about 12 times the mass of Charon. Knowing the diameters and masses of Pluto and Charon

allows astronomers to calculate that their densities are both about 2 g/cm^3. This indicates that Pluto and Charon must contain about 35 percent ice and 65 percent rock.

The best photos by the Hubble Space Telescope reveal almost no surface detail, but you know enough about icy moons to guess that Pluto has craters and probably shows signs of tidal heating caused by interaction with its large moon Charon. The New Horizons spacecraft will fly past Pluto in July 2015, and the images radioed back to Earth will certainly show that Pluto has some surprising features.

What Defines a Planet?

To understand why Pluto is no longer considered a planet, you should recall what you have already learned about the Kuiper belt (Chapter 8). Since 1992 astronomers have discovered more than a thousand icy bodies orbiting beyond Neptune. There may be as many as 100 million Kuiper belt objects larger than 1 km in diameter. They are understood to be icy planetesimals left over from the outer solar nebula. Some of the KBOs are quite large, and one, Eris, has about the same diameter as Pluto but is 27 percent more massive. Three other KBOs found so far, Sedna, Quaoar (pronounced *kwah-o-wahr*), and Orcus, are half the size of Pluto or larger. Eris, Quaoar, and Orcus are known to have moons of their own. In that way, they resemble Pluto and its family of moons.

A bit of comparative planetology shows that Pluto is not related to the Jovian or Terrestrial planets; it is obviously a member of a newfound family of worlds that orbit beyond Neptune. These bodies must have formed at about the same time as the eight classical planets of the solar system, but they did not grow massive enough to clear their orbital zones of remnant planetesimals and consequently remain embedded among a swarm of other objects in the Kuiper belt.

The IAU's criteria for full planet status is that an object must be large enough that its gravity has pulled it into a spherical shape and also large enough to dominate and gravitationally clear its orbital region of most or all other objects over a span of billions of years. Eris and Pluto, the largest objects found so far in the Kuiper belt, and Ceres, the largest object in the asteroid belt, are too small to clear their orbital zones of other objects and therefore do not meet the standard for being called planets. On the other hand all three are large enough to be spherical, so they are the prototypes of a new class of objects defined by the IAU as **dwarf planets.**

Pluto and the Plutinos

No, this section is not about a 1950s rock band. It is about the history of the outer solar system, and it will take you back 4.6 billion years to watch the planets form.

More than a hundred of the Kuiper belt objects are caught with Pluto in a 3:2 resonance with Neptune. That is, they orbit the sun twice while Neptune orbits three times. This subset of KBOs have been named **plutinos.** The plutinos formed in the outer solar nebula, but how did they get caught in resonances with Neptune? As you learned earlier, some models of the formation of the planets suggest that Uranus and Neptune may have migrated outward early in the solar system's history. As Neptune moved farther from the sun, its orbital resonances could have swept up small objects like a kind of planetary snow plow. The plutinos are caught in the 3:2 resonance, and other KBOs are caught in other resonances. The evidence appears to support those models which predict that Uranus and Neptune migrated outward.

The migration of the outer planets would have dramatically upset the motion of some of these Kuiper belt objects, and some could have been thrown inward where they could interact with the Jovian planets. Some of those objects may have been captured as moons, and astronomers wonder if moons such as Neptune's Triton could have started life as KBOs. Other objects may have hit bodies in the inner solar system and caused the late heavy bombardment episode especially evident on the surface of Earth's moon (look back to Chapter 9). The small frozen worlds on the fringes of the solar system may hold clues to the formation of the planets 4.6 billion years ago and the subsequent history of Earth.

SCIENTIFIC ARGUMENT

What evidence indicates that cataclysmic impacts have occurred in our solar system?

To build this argument, you can cite plenty of evidence. The retrograde and highly eccentric orbits, respectively, of Neptune's moons Triton and Nereid, the peculiar rotation of Uranus, and the size of Pluto's satellite Charon all hint that impacts and encounters with large planetesimals have been important in the history of these worlds. Furthermore, the existence of planetary rings suggests that impacts have scattered small particles and replenished the ring systems. Even in the inner solar system, the retrograde rotation of Venus, the high density of Mercury, the smooth northern lowlands of Mars, and the formation of Earth's moon are possible consequences of major impacts when the solar system was young.

Now assemble evidence in a new argument. **How does the origin of the plutinos give a clue about one possible source of impacting bodies in the solar system?**

What Are We? Trapped

No one has ever been further from Earth than the moon. We humans have sent robotic spacecraft to visit most of the larger worlds in our solar system, and we have found them strange and wonderful places, but no human has ever set foot on any of them. We are trapped on Earth.

We lack the technology to leave Earth. Getting away from Earth's gravitational field is difficult and calls for very large rockets. America built such rockets in the 1960s and early 1970s. They could send astronauts to the moon, but such rockets no longer exist. The best technology today can carry astronauts just a few hundred kilometers above Earth's surface to orbit above the atmosphere. The United States and other nations are considering sending humans back to the moon and eventually to Mars, but budget limitations have delayed specific plans. Does Earth's civilization have the resources to build spacecraft capable of carrying human explorers to other worlds? We'll have to wait and see.

In the previous chapter, you discovered another reason we Earthlings are trapped on Earth. We have evolved to fit the environment on Earth. None of the planets or moons you explored in this chapter would welcome you. Lack of air, and extreme heat or cold, are obvious problems, but, also, Earthlings have evolved to live with Earth's gravity. Astronauts in space for just a few weeks suffer biomedical problems because they are no longer in Earth's gravity. Living in a colony on Mars or the moon might raise similar problems. Just getting to the outer planets would take decades of space travel; living for years in a colony on one of the Jovian moons under low gravity and exposed to the planet's radiation belts may be beyond the capability of the human body. We may be trapped on Earth not because we lack large enough rockets but because we need Earth's protection.

It seems likely that we need Earth more than it needs us. The human race is changing the world we live on at a startling pace, and some of those changes could make Earth less hospitable to human life. All of your exploring of un-Earthly worlds serves to remind you of the nurturing beauty of our home planet.

Study and Review

Summary

- The Jovian planets—Jupiter, Saturn, Uranus, and Neptune—are large, massive low-density worlds located in the outer solar system.
- Atmospheres of the Jovian planets are marked by **belt–zone circulation (p. 211)** that produces cloud belts parallel to their equators.
- Models of planetary interiors can be calculated based on each planet's density and **oblateness (p. 211)**, the fraction by which its equatorial diameter exceeds its polar diameter.
- Jupiter and Saturn are composed mostly of liquid metallic hydrogen, and for this reason they are sometimes referred to as "liquid giants."
- Uranus and Neptune contain abundant water in solid form and are sometimes called "ice giants."
- All the Jovian planets have large systems of satellites and rings that have had complex histories. **Regular satellites (p. 211)** are generally larger, are close to the parent planet, revolve in the **prograde (p. 211)** direction, and have low orbital inclinations; **irregular satellites (p. 211)** are generally small, are far from the parent planet, and have high orbital inclinations.
- Jupiter is observed to have heat flowing out of it at a high rate, indicating that its interior is very hot.
- Jupiter has a core of heavy elements surrounded by a deep mantle of **liquid metallic hydrogen (p. 211)** in which the planet's large and strong magnetic field is generated.
- The **magnetosphere (p. 211)** around Jupiter traps high-energy particles from the sun to form intense radiation belts.
- Jupiter's atmosphere contains three layers of clouds formed of hydrogen-rich molecules, including water and ammonia. The cloud layers are at altitudes corresponding to the condensation temperatures of their respective main constituents.
- The cloud stripes parallel to Jupiter's equator consist of light-colored zones that are high-pressure regions of rising gas plus darker belts that are lower-pressure areas of sinking gas.
- Spots in Jupiter's atmosphere, including the Great Red Spot, are circulating weather patterns.
- The four large **Galilean moons (p. 213)** show signs of geologic activity. **Grooved terrain (p. 213)** on Ganymede, smooth ice and cracks on Europa, and active volcanoes on Io show that that **tidal heating (p. 216)** driven by **orbital resonances (p. 216)** has made these moons active.
- Many of Jupiter's moons are small, rocky bodies that are probably captured objects. They are too small to retain heat and are not geologically active.
- Jupiter's rings are composed of dark particles that are bright when illuminated from behind, which is called **forward scattering (p. 216).** This shows that the particles are very small and are probably dust from meteoroid impacts on moons.
- Jupiter's rings, like all of the rings in the solar system, lie inside the planet's **Roche limit (p. 217),** within which tidal stress can destroy a moon or prevent one from forming.
- Saturn is less dense than water and contains a small core and less metallic hydrogen than Jupiter.
- Cloud layers on Saturn occur at the same temperature as those on Jupiter, but because Saturn is further from the sun and colder the cloud layers are deeper in the hydrogen atmosphere below a layer of methane haze and so are less prominent.
- Saturn's moons are icy and mostly heavily cratered.
- Saturn's largest moon, Titan, has a cold, cloudy nitrogen atmosphere. It may have had methane falling as rain on its icy surface and forming lakes and rivers.
- Sunlight entering Titan's atmosphere can convert methane into complex carbon-rich molecules to form haze and particles that settle out to coat the surface with dark organic material.
- Enceladus has a light surface with some uncratered regions. Geysers of water and ice containing organic molecules vent from the region around its south pole and provide ice particles to the extensive low-density E ring.
- Saturn's rings are composed of icy particles ranging in size from boulders to dust. The composition and brightness of the ring particles vary from place to place in the rings.
- Grooves in the rings can be produced by orbital resonances, or waves that propagate through the rings, caused by moons near or within the rings.
- Narrow rings as well as sharp ring edges can be produced by **shepherd satellites (p. 225).**
- The material observed now in the Jovian planets' rings cannot have lasted since the formation of the solar system. The rings are understood to be replenished occasionally with material produced by meteoroids, asteroids and comets colliding with moons.
- Uranus is much less massive than Jupiter, and its internal pressure cannot produce liquid hydrogen. Models indicate it has a heavy-element core and a mantle of solid or slushy ice and rocky material below a hydrogen-rich atmosphere. Little heat flows out of Uranus, so it cannot be very hot inside.
- Uranus's atmosphere is almost featureless at visual wavelengths with a pale blue color caused by traces of methane, which absorbs red light. Images at selected wavelengths can be enhanced to show traces of belt–zone circulation.
- Uranus rotates on its side, perhaps because of a major impact or tidal interactions with other planets during its early history.
- The larger moons of Uranus are icy and heavily cratered, with signs on some of past geological activity, including **ovoids (p. 223)** on Miranda.
- The rings of Uranus, discovered by stellar **occultations (p. 226),** are narrow hoops confined by shepherd satellites. The particles appear to be ice with traces of methane darkened by the radiation belt.
- Neptune, like Uranus, is an ice giant with no liquid hydrogen. Unlike Uranus, Neptune does have heat flowing outward from its interior.
- The atmosphere of Neptune, marked by traces of belt–zone circulation, is rich in hydrogen and colored blue by traces of methane.
- Neptune's satellite system is odd in that distant Nereid follows an elliptical orbit and Triton orbits backward. These may be signs of catastrophic encounters with other objects early in the solar system's history.
- Triton is icy with a thin atmosphere and frosty polar caps. Smooth areas suggest past geological activity, and dark smudges mark the location of active nitrogen geysers.
- Neptune's rings are made of icy particles in narrow hoops and contain arcs produced by the gravitational influence of one or more moons.
- Pluto is a small world with five known moons, one of which, Charon, is quite large in relation to Pluto. The moons' orbital plane and Pluto's equator are highly inclined to Pluto's orbit around the sun. Pluto's composition is mostly rock with a substantial amount of ice.
- Pluto is a member of a family of Kuiper belt objects orbiting beyond Neptune. At least one of those objects, Eris, is a bit larger than Pluto. The **International Astronomical Union (IAU) (p. 223)** that decides definitions and criteria for naming celestial objects redefined Pluto in 2006 as a **dwarf planet (p. 234).**
- Some Kuiper belt objects called **plutinos (p. 234)** follow orbits like Pluto that have an orbital resonance with Neptune.
- Some model calculations suggest that Uranus and Neptune formed closer to the sun and migrated outward, pushing millions of icy bodies into orbital resonances farther from the sun to form the Kuiper belt. Pluto may be one of those objects.

Study and Review

Review Questions

1. Why is Jupiter so much richer in hydrogen and helium than Earth?
2. How can Jupiter have a liquid interior and not have a definite liquid surface?
3. How does the dynamo effect account for the magnetic fields of Jupiter, Saturn, Uranus, and Neptune?
4. Why are the belts and zones on Saturn less distinct than those on Jupiter?
5. Why do astronomers conclude that none of the Jovian planets' rings can be left over from the formation of the planets?
6. How can a moon produce a gap in a planetary ring system?
7. Explain why the amount of geological activity on Jupiter's moons varies with distance from the planet.
8. What makes Saturn's F ring and the rings of Uranus and Neptune so narrow?
9. Why is the atmospheric activity of Uranus less than that of Saturn and Neptune?
10. Why do astronomers suspect that Saturn's moon Enceladus is geologically active?
11. What are the seasons on Uranus like?
12. Why are Uranus and Neptune respectively green-blue and blue?
13. What evidence is there that Neptune's moon Triton has been geologically active recently?
14. How do astronomers account for the origin of Pluto?
15. What evidence indicates that catastrophic impacts have occurred in the solar system's past?
16. **How Do We Know?** Why would you expect research in archaeology to be less well funded than research in chemistry?

Discussion Questions

1. Some astronomers argue that Jupiter and Saturn are unusual, while other astronomers argue that all planetary systems should contain one or two such giant planets. What do you think? Support your argument with evidence.
2. Why don't the Terrestrial planets have rings? If you were to search for a ring among the Terrestrial planets, where would you look first?

Problems

1. What is the maximum angular diameter of Jupiter as seen from Earth? Repeat this calculation for Neptune. Useful data can be found in Celestial Profiles 7 and 10. (*Hint:* Use the small-angle formula, Chapter 3.)
2. What is the angular diameter of Jupiter as seen from Callisto? From Amalthea? Useful data can be found in Appendix Table A-11. (*Hint:* Use the small-angle formula, Chapter 3.)
3. Measure the polar and equatorial diameters of Saturn in the photograph in Celestial Profile 8 and calculate the planet's oblateness using the definition given in this chapter.
4. If you observe light reflected from Saturn's rings, you should see a redshift at one edge of the rings and a blueshift at the other edge. If you observe a spectral line and see a difference in wavelength of 0.0560 nm between opposite edges of the rings, and the unshifted wavelength (observed in the laboratory) is 500. nm, what is the orbital velocity of particles at the outer edge of the rings? (*Hint:* Use the Doppler shift formula, Chapter 6.)
5. One way to recognize a distant planet is by its motion along its orbit. If Uranus circles the sun in 84 years, how many arc seconds will it move in 24 hours? (For the purposes of this problem, ignore the motion of Earth.)
6. If Uranus's ε ring is 50 km wide and the orbital velocity of Uranus is 6.8 km/s, how long should the occultation last that you expect to observe from Earth when the ring crosses in front of the star? (For the purposes of this problem, ignore the motion of Earth.)
7. If Neptune's clouds have a temperature of 60 K, at what wavelength will they radiate the most energy? (*Hint:* Use the Wien's law formula, Chapter 6.)
8. How long did it take radio commands to travel from Earth to *Voyager 2* as it passed Neptune? Relevant data can be found in Appendix Table A-10. Assume Earth and Neptune were as close as possible during the *Voyager 2* encounter.
9. What is the angular diameter of Pluto as seen from the surface of Charon? Relevant data can be found in Figure 11-16.
10. The orbital period of Charon is given in the chapter text. Given the average distance of Charon from Pluto is 19,600 km, calculate the mass of Pluto. Assume the mass of Charon is negligible compared with the mass of Pluto. (*Hints:* Express the orbital radius in meters and the period in seconds. Then, use the circular orbit velocity formula, Chapter 4.)

Learning to Look

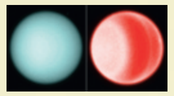
1. This photo shows a segment of the surface of Jupiter's moon Callisto. Why is the surface mostly dark? Why are some craters dark and some white? What does this image tell you about the history of Callisto?

NASA/JPL/University of Arizona

2. The *Cassini* spacecraft recorded this photo of Saturn's A ring and Encke's division. What do you see in this photo that tells you about processes that confine and shape planetary rings?

NASA/JPL/Space Science Institute

3. Two images are shown here of Uranus's northern hemisphere, one as it would look to the eye and the other through a red filter that enhances methane clouds. What do the atmospheric features tell you about circulation on Uranus?



NASA/JPL/STScI

Great Debates

1. *New Home for Earthlings.* The sun will eventually die; during its death, it could expand to the orbit of Earth. Humans will not survive this expansion of the sun. Two good new home choices for Earthlings are Enceledus and Titan. On which of these objects should NASA spend our tax dollars to study such questions as developing technology to transport humans to this new location?
   a. Use at least three vocabulary words and data from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
2. *Comet Impact.* As discussed in the text, comet Shoemaker-Levy 9 (SL9) struck Jupiter in 1994. Should Earthlings be worried about comets like SL9 striking the Earth? Why or why not?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional information that supports your claim, including details of past impacts to Earth.
   c. Cite your sources.
3. *Galilean Moons?* The year is 2200, and the world has become overrun with humans. However, we now have the technology to colonize the Galilean moons. You have been appointed to a committee to handle the overcrowding situation on Earth. You decide to sell real estate on these moons. There need to be three kinds of colonies created: (1) new correctional facilities for the government, (2) luxury and vacation homes for the wealthy, (3) middle-class housing, like the suburbs on Earth. Which Galilean moon is best suited for which colony, and why?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
4. *Pluto's Fate.* The International Astronomical Union (IAU) called for a vote on the definition of *planet,* and thus Pluto was reclassified as a dwarf planet. The IAU invited their members for the vote. Should the IAU have considered the opinion of the general public? Do you feel that the outcome of the vote is correct? Should Pluto be classified as a planet?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
5. *Planet X.* The wealthy Percival Lowell helped build the Lowell Observatory in 1894. Lowell was convinced a Planet X was involved in the perturbing of Uranus's orbit so that Uranus passed inside the orbit of Neptune. Pluto was thought to be the Planet X, but since its downgrade to a dwarf planet, the search continues for Planet X. Do you believe an undiscovered Planet X still exists that alters the orbit of Uranus? Should tax dollars be spent to find Planet X?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

ENHANCED WebAssign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Roche Limit
- Planetary Atmospheres
- Magnetic Fields of Uranus
- Uranus's Ring Detection

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 7: Planetary Atmospheres and Their Retention

One of the textbook authors happened to be at an astronomy research conference when detection of the first extrasolar planet orbiting a solar-type star, 51 Pegasi, was announced in 1995. As you read in a previous chapter, that planet has a mass in the Jovian range, with a lower limit of about half the mass of Jupiter (more massive than Saturn). The additional surprising information for the astronomers at the conference, and in fact all over the world, was that 51 Pegasi's planet has an orbital radius of 0.05 AU and thus an estimated surface temperature (depending on its reflectivity) of at least 1200 K.

Wait, we said. How can that be? Wouldn't a Jovian planet, made mostly of light gasses like hydrogen and helium, vaporize at that temperature? An auditorium full of astronomers proceeded to ignore the speaker, pulled paper, pencils, and calculators out of backpacks, and raced to estimate how long it would take for Saturn to boil away at 1200 K. And the answer is: It wouldn't, not even over billions of years. Planet formation theorists continue to maintain that a Jovian-mass planet can form by condensation and accretion only beyond a protoplanetary nebula's "snow line," where hydrogen compounds such as water, ammonia, and methane can condense in space as solid particles. However, it seems that if a Jovian planet forms where it "should" and then somehow migrates closer to the stellar furnace, it would have enough gravity to hold onto its material, even at temperatures of thousands of degrees, for the lifetime of a planetary system.

You can verify this for yourself semiqualitatively. In Section 4 of Virtual Astronomy Lab 7, "Planetary Atmospheres and Their Retention," there is a plot showing average speeds as a function of temperature for various atoms and molecules found in planetary atmospheres. On that plot there are also data points representing each planet by its upper atmosphere temperature and escape velocity. (Essentially the same plot is on page 198 of the textbook.) Make a new version of that plot by extending the *x*-axis and the curve representing the speed of hydrogen molecules (H_2) to the right, to a temperature of 1200 K. Then, replot the data point for Saturn, currently in the upper left corner, at the same vertical axis position (that is, keeping the planet's escape velocity constant) but at a new horizontal axis position of 1200 K. You will find a very large vertical distance between the speed of hydrogen molecules in the atmosphere of a "hot Saturn" and Saturn's escape velocity. What does that mean? Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 6: Tides and Tidal Forces

Tides are everywhere, not only at the ocean's shore, and not only on Earth. Newton's understanding of gravity implies that ocean tides are just the most familiar manifestation of a universal phenomenon. In general, the gravitational interaction of any two objects will stress and stretch them because of the difference between the gravitational forces exerted at opposite sides of the objects. This is what an astronomer means by "tides," not the rising and falling of the ocean during the day, but the ultimate cause, the moon's gravity stretching Earth. At the same time, the tidal force due to Earth's gravity is stretching the moon.

Imagine that two objects interacting gravitationally come so close together that the tidal forces are comparable to their material strengths. They might break apart. The maximum separation at which this catastrophe can happen is called the Roche limit (or Roche radius), named after the French astronomer who first calculated it.

In this chapter you learned that the rings of Jupiter, Saturn, Uranus, and Neptune are located within the Roche limits of their respective planets. Does this mean that the Jovian planets' rings are remnants of moons or other objects that strayed too close and were destroyed by tidal forces? Not necessarily. Another way to define the Roche limit is the radius within which tidal forces prevent existing particles from combining. Material chipped from Jovian moons by meteorite impacts may be the source of most of the rings. Such material in orbit within a planet's Roche radius cannot recombine into larger objects.

In another chapter you will see images of pieces of Comet Shoemaker-Levy 9, which was shredded by tidal stress as it passed by Jupiter two years before returning to collide with the planet. Also, chains of small craters on Earth's moon and on Jupiter's moon Callisto appear to be the result of impacts by trains of comet fragments like Shoemaker-Levy 9. Perhaps such tidal encounters are fairly common in the life of the solar system.

Section 3 of Virtual Astronomy Lab 6, "Tides and Tidal Forces," guides you through calculating Roche limits for planets and moons of various masses and sizes.

Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

12

Meteorites, Asteroids, and Comets

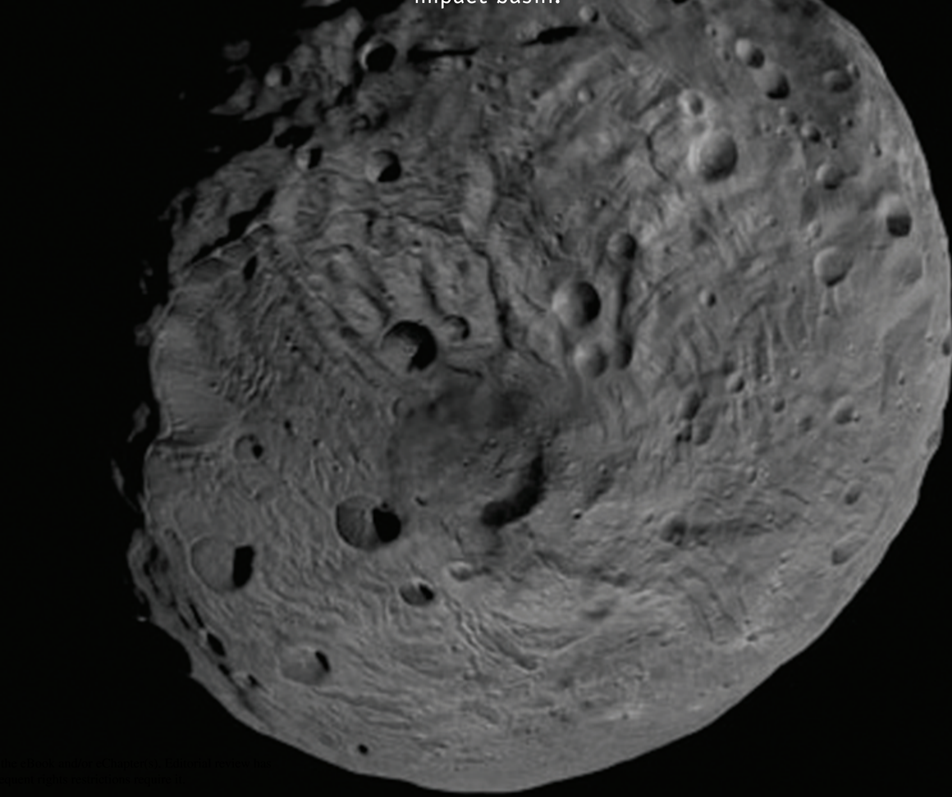
Guidepost

In Chapter 8 you began your study of planetary astronomy by considering evidence about how our solar system formed. In the three chapters that followed you surveyed the planets and found more clues about the origin of the solar system but also learned that most traces of the early histories of the planets have been erased by geological activity or other processes. Now you can study smaller, less-altered objects that tell more about the era of planet building.

Asteroids and comets are unevolved objects, leftover planet construction "bricks." You will find them much as they were when they formed 4.6 billion years ago. Meteors and meteorites are fragments of comets and asteroids that arrive at Earth and can give you a close look at those ancient planetesimals. As you explore, you will find answers to four important questions:

- **Where do meteors and meteorites come from?**
- **What are asteroids?**
- **What are comets?**
- **What happens when asteroids and comets hit Earth and other planets?**

As you reach the end of this twelfth chapter, you will have acquired real insight into your place in nature. You live on the surface of a planet. Are any other planets inhabited? That is the subject of Chapter 20.



This image of the asteroid Vesta, made from a distance of 2700 km (1700 mi) by NASA's Dawn spacecraft, shows Vesta's south polar region, which is dominated by the Rheasilvia impact basin.

Visual-wavelength image

NASA/JPL-Caltech/UCLA/MPS/DLR/IDA

NASA/JPL-Caltech/UCLA/MPS/DLR/IDA

When they shall cry “PEACE, PEACE”
then cometh sudden destruction!
COMET’S CHAOS?—
What Terrible events will the Comet bring?

FROM A PAMPHLET PREDICTING THE END OF THE WORLD BECAUSE OF THE APPEARANCE OF COMET KOHOUTEK IN 1973

One afternoon in 1954, while Mrs. E. Hulitt Hodges of Sylacauga, Alabama, lay napping on her living room couch, an explosion and a sharp pain jolted her awake. Analysis of the brick-sized rock that smashed through the ceiling and bruised her left leg showed that it was a meteorite. Mrs. Hodges is the only person known to have been injured by a meteorite. Coincidentally, she lived right across the street from the Comet Drive-In Theater.

Meteorites arrive from space all over the Earth every day, although not as spectacularly as the one that struck Mrs. Hodges. You will learn in this chapter that meteorites are fragments of asteroids, and that asteroids, as well as their icy cousins the comets, carry precious clues about conditions in the solar nebula from which the sun and planets formed. Because you cannot easily visit comets and asteroids, you can begin by learning about the pieces of those bodies that come to you.

12-1 Meteoroids, Meteors, and Meteorites

You learned some things about meteorites in Chapter 8 when you studied evidence for the age of the solar system. There you saw that the solar system includes small particles called *meteoroids.* Some of them collide with Earth’s atmosphere at speeds of 10 to 70 km/s. Friction with the air heats the meteoroids enough so that they glow, and you see them vaporize as streaks across the night sky. Those streaks are called *meteors* (“shooting stars”).

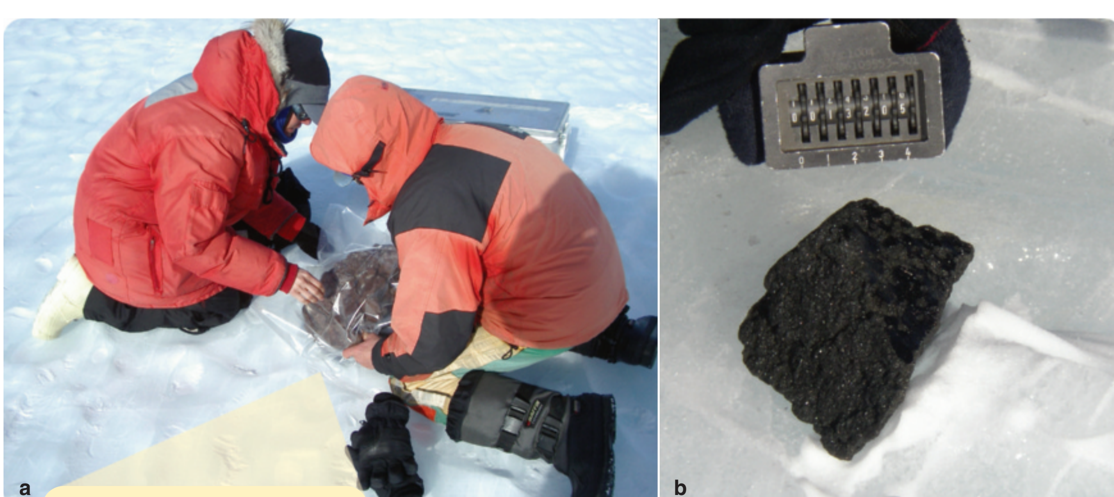
If a meteoroid is big enough and holds together well enough, it can survive its plunge through the atmosphere and reach Earth’s surface. Once the object strikes Earth’s surface, it is called a *meteorite* (“-ite” being the Greek root for “rock”). As you will learn later in this chapter, the largest of those objects can blast out craters on Earth’s surface, but such big impacts are extremely rare. The great majority of meteorites are too small to form craters.

What can meteorites and meteors tell you about the origin of the solar system? To answer that question, you can consider their compositions and their orbits.

Composition of Meteorites

One of the best places to look for meteorites turns out to be certain parts of Antarctica. The nearest Earth rocks are buried under miles of ice, so any rock you find there must have fallen from space. The slow flow of the Antarctic ice cap toward the ocean concentrates meteorites in areas where the moving ice runs into mountain barriers, slows down, and evaporates. Teams of scientists travel to Antarctica and ride snowmobiles in systematic sweeps across the ice each Southern Hemisphere summer to rec over meteorites (■ Figure 12-1). After 25 years of work, more than half of the 40,000 meteorites in human hands are from Antarctica.

Meteorites can be divided into three broad composition categories. **Iron meteorites** (■ Figure 12-2a) are solid chunks of iron and nickel. **Stony-iron meteorites** (Figure 12-2c) are mixtures of iron and stone. **Stony meteorites** (Figure 12-2d)



Large or small, meteorites are sealed airtight and refrigerated until they can be studied.

© Cengage Learning 2014; Images: ANSMET/NSF/NASA; courtesy Monika Kress

■ **Figure 12-1**

Braving bitter cold and high winds, teams of scientists riding snowmobiles search for meteorites that fell long ago in Antarctica and are exposed as the ice evaporates. When a meteorite is found, it is photographed where it lies, assigned a number, and placed in an airtight bag. Thousands of meteorites have been collected in this way, including rare meteorites from the moon and Mars.

Figure 12-2

The main types of meteorites have distinctive characteristics. (a) An iron meteorite, actually composed mostly of a mixture of iron and nickel. (b) Sliced, polished, and etched with acid, iron meteorites show what is called a Widmanstätten pattern of large crystals, indicating that this material cooled very slowly from a molten state and must have been inside a fairly large object. (c) A stony-iron meteorite. (d) A stony meteorite. (e) Carbonaceous chondrite meteorites are a rare type of stony meteorite that is rich in carbon, making the rock very dark.

are silicate masses that resemble Earth rocks. **Carbonaceous chondrites** (pronounced *kon-drite;* Figure 12-2e) are a special type of stony meteorite.

Iron meteorites make up only 6 percent of meteorites witnessed falling from the sky, so you can estimate that only about 6 percent of meteoroids traveling through space near Earth have iron composition. Then you might wonder, why do iron meteorites make up 66 percent of meteorites found on the ground? Because iron meteorites don't look like ordinary rocks; they are easy to recognize because they are heavy, dense lumps of metal. If you trip over one on a hike, you are more likely to recognize it as something odd, carry it home, and show it to the local museum. Also, some stony meteorites deteriorate rapidly when exposed to weather; irons are made of stronger material and generally survive longer. The recognizability and durability of iron meteorites means there is a **selection effect** that makes it more likely they will be found than other types of meteorites (**How Do We Know? 12-1**).

When iron meteorites are sliced open, polished, and etched with acid, they reveal regular bands called **Widmanstätten patterns** (pronounced *VEED-mahn-state-en*) (Figure 12-2b) caused by alloys of iron and nickel that formed crystals as the molten metal cooled and solidified. The size and shape of the bands indicate that the molten metal cooled very slowly, indicating

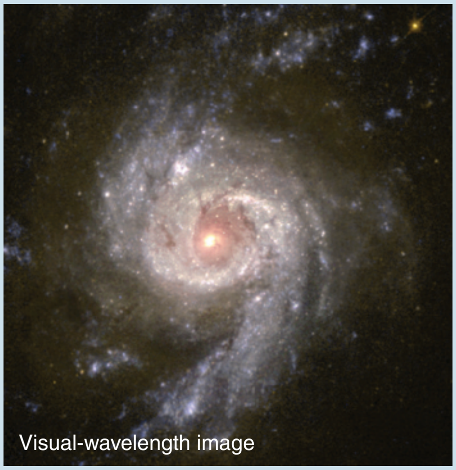
How Do We Know? 12-1

Selection Effects

How might scientists go astray if they are careless in selecting subjects for study? Scientists must plan ahead and design their research projects with great care. Biologists studying insects in the rain forest, for example, must choose which ones to catch. They can't catch every insect they see, so they might decide to catch and study any insect that is red. If they are not careful, a selection effect could bias their data and lead them to incorrect conclusions without their ever knowing it.

For example, suppose you needed to measure the speed of cars on a highway. There are too many cars to measure every one, so you might reduce the workload and measure only red cars. It is quite possible that this selection criterion will mislead you because people who buy red cars may be more likely to be younger and drive faster. Should you instead measure only brown cars? No, because older, more sedate people might tend to buy brown cars. Only by very carefully designing your experiment can you be certain that the cars you measure are traveling at typical speeds.

Astronomers understand that what you see through a telescope depends on what you notice, and that is powerfully influenced by what are called "selection effects." The biologists in the rain forest, for example, should not catch and study only red insects. Often, the most brightly colored insects are poisonous or at least taste bad to predators. Catching only red insects could produce a result highly biased by a selection effect.



Things that are bright and beautiful, such as spiral galaxies, may attract a disproportionate amount of attention. Scientists must be aware of such selection effects.

a location inside bodies at least 30 km in diameter. (In comparison, a small lump of molten metal exposed in space would cool in just a few hours.) On the other hand, the iron meteorites do not show effects of the very high pressures that would exist deep inside a planet. Evidently, iron meteorite material formed in the cooling interiors of planetesimal-sized objects, smaller than planets. This is one important clue to the origin of meteorites.

In contrast to irons and stony-irons, stony meteorites (Figure 12-2d) are common among meteorites seen falling to Earth. Although there are many different types of stony meteorites, you can classify them into two main categories depending on their physical properties and chemical content: **chondrites** and **achondrites.**

Chondrites look like dark gray, granular rocks (Figure 12-2e). In general, chondrites contain some volatiles including water and organic (carbon) compounds. A few chondrites evidently formed in the presence of liquid water. Most types of chondrites also contain **chondrules,** small round bits of glassy rock only a few millimeters across. To be glassy rather than crystalline, the chondrules must have cooled from a molten state quickly, within a few hours. One hypothesis is that chondrules are bits of matter from the inner part of the solar nebula, near the sun, that were blown outward by winds or jets to cooler parts of the nebula where they condensed and were later incorporated into larger rocks. Another hypothesis is that the chondrules were once solid bits of matter that were melted by shock waves spreading through the solar nebula and then resolidified.

Among the chondrites, the carbonaceous chondrites are a rare but quite important subtype. These dark gray, rocky meteorites are especially rich in water, other volatiles, and organic compounds. Those substances all would have been lost if the meteoroid had been heated even to room temperature. Many carbonaceous chondrites also contain small, irregular inclusions rich in calcium, aluminum, and titanium. Called **CAIs,** for calcium-aluminum-rich inclusions, these bits of matter are highly refractory; that is, they vaporize or condense only at very high temperatures. If you could scoop out a portion of the sun's photosphere and cool it, the first particles to solidify would have the chemical composition of CAIs. As the temperature fell, other materials would become solid in accord with the condensation sequence described in Chapter 8.

When the cooling blob of solar-composition material finally reached room temperature, you would find that almost all of the hydrogen, helium, and some other gases such as argon and neon had escaped and that the remaining lump would have almost exactly the same overall chemical composition as the main mass of a carbonaceous chondrite. Carbonaceous chondrite meteorites are evidently very old samples of the solar nebula, confirmed by the fact that CAIs have radioactive ages equal to the oldest of any other solar system material.

The chondrites show properties ranging from carbonaceous chondrites, most of which have avoided being heated or modified, to other chondrites in which the material was slightly heated and somewhat altered from the form in which it first solidified.

Chondrites in general offer us the best direct information about conditions and processes occurring in the earliest days of the solar nebula when planetesimals and planets were forming.

Stony meteorites called achondrites contain no chondrules and also lack volatiles. These rocks appear to have been subjected to intense heat that melted the chondrules and completely drove off the volatiles, resulting in a composition similar to Earth's basalts.

The different types of meteorites evidently had a wide variety of histories. Carbonaceous chondrites are like unaltered lumps of condensed solar nebula; achondrites seem like pieces of lava flows, whereas stony-iron and iron meteorites apparently were once deep inside the molten interiors of differentiated objects. Meteorites provide evidence that the young solar system was a complicated place.

Orbits and Origins of Meteoroids

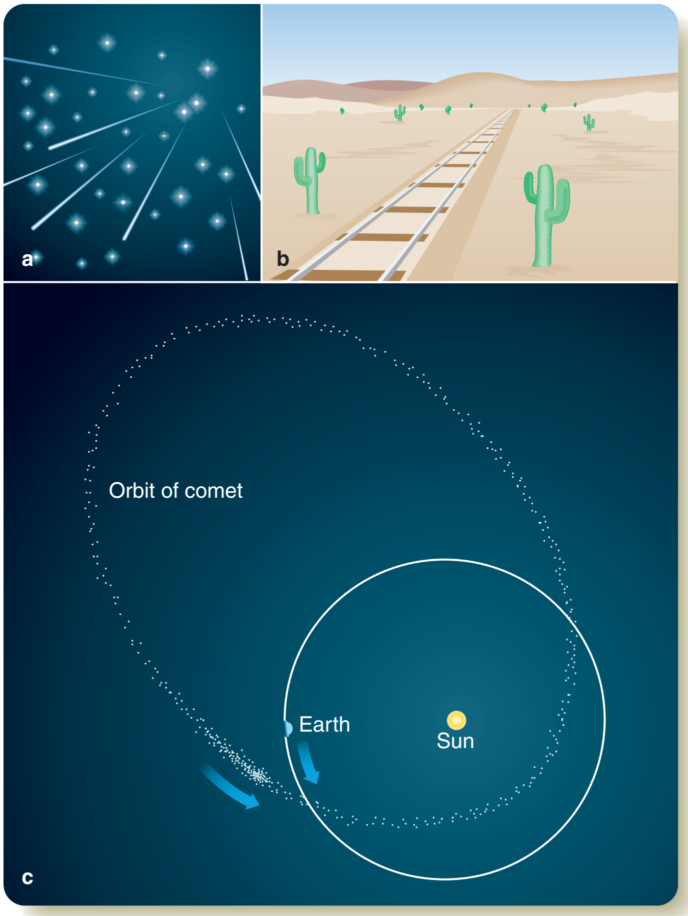
A typical meteoroid has roughly the mass of a paper clip and vaporizes at an altitude of about 80 km (50 mi) above Earth's surface. The meteor trail points back along the path of the meteoroid, so if you study the direction and speed of meteors, you can get clues to their orbits in the solar system before they encountered Earth.

One way to backtrack meteor trails is to observe **meteor showers.** On any clear night, you can see 3 to 15 meteors per hour, but on some nights you can see a shower of hundreds of meteors per hour that are obviously related to each other, seeming to come from a single area of the sky, called the **radiant** of the shower (■ Figure 12-3a). Meteor showers are named after the constellation from which they seem to radiate; for example, the Perseid shower seen in mid-August radiates from the constellation Perseus. The fact that the meteors in a shower appear to come from a single point in the sky means that the meteoroids were actually traveling through space along parallel paths. You see their fiery tracks through Earth's atmosphere in perspective, so they appear to come from a single radiant point, just as parallel railroad tracks seem to come from a single point on the horizon (Figure 12-3b).

The meteoroids in a meteor shower are understood to be dust and debris released from the icy nucleus of a comet. When Earth passes through this stream of material, you see a meteor shower (Figure 12-3c). In some cases the comet has wasted away and is no longer visible, but in other cases the comet is still prominent, although located somewhere else along its orbit. For example, each May, Earth comes near the orbit of Comet Halley, and you can see the Eta Aquarid shower. Each October, Earth passes near the other side of the orbit of Comet Halley, and the Orionid shower appears.

Even when there is no shower, you still can occasionally see meteors that are called **sporadic meteors** because they are not part of specific showers. Studies of the trails of sporadic meteors reveal that they have a dual source: Many come from comets, as do meteors in showers, but a few come from the asteroid belt. Objects that are big and durable enough to become meteorites on the ground seem always to come from the asteroid belt.

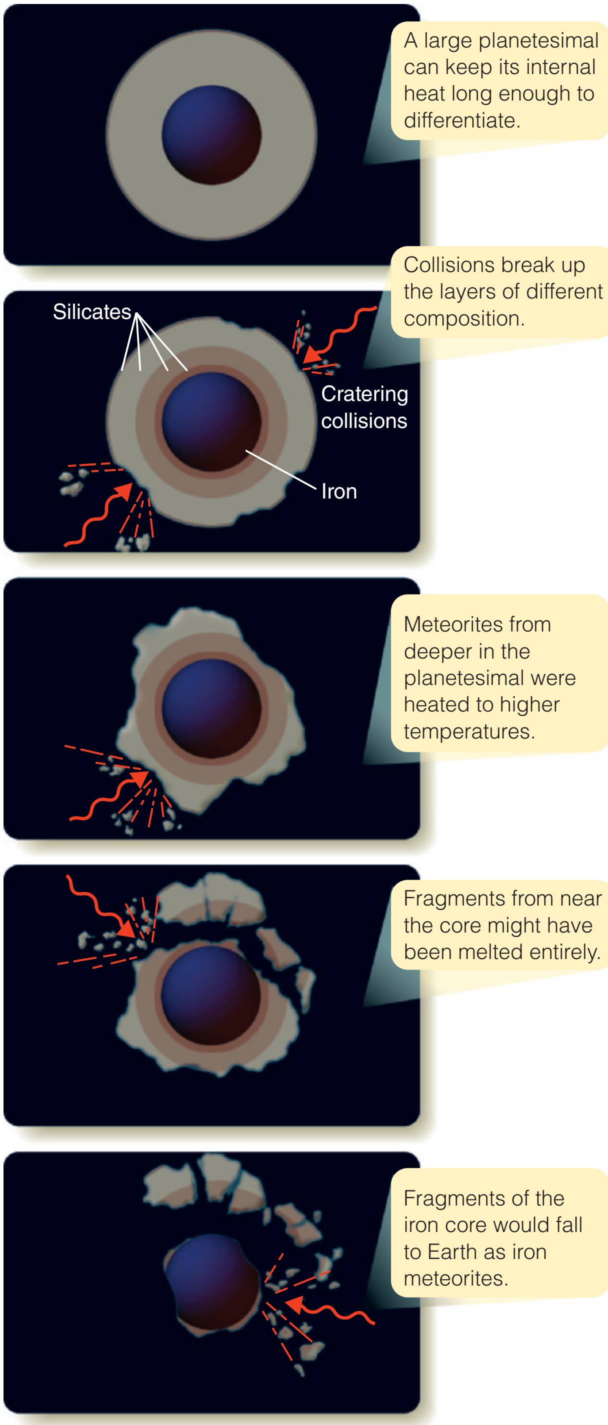
The physical characteristics of meteorites provide evidence that they are fragments of parent bodies in the asteroid belt that were large enough to grow hot from radioactive decay or other processes. They then melted and differentiated to form iron-nickel cores and rocky mantles. The molten iron cores would have been well insulated by the thick rocky mantles, so that the iron would have cooled slowly enough to produce big crystals that result in Widmanstätten patterns. Some stony meteorites that have been strongly heated appear instead to have come from the mantles or surfaces of such bodies. Stony-iron meteorites apparently come from boundaries between stony mantles and iron cores. Collisions could break up such differentiated bodies



■ **Figure 12-3**

(a) Meteors in a meteor shower enter Earth's atmosphere along parallel paths, but perspective makes them appear to diverge from a single point in the sky. (b) Similarly, parallel railroad tracks appear to diverge from a point on the horizon. (c) Meteors in a shower are debris left behind as a comet's icy nucleus vaporizes. The rocky and metallic bits of matter spread along the comet's orbit. If Earth passes through such material, you can see a meteor shower.

The Origin of Meteorites



© Cengage Learning 2014; Adapted from a diagram by Clark Chapman

Figure 12-4

Some of the planetesimals that formed early in the solar system's history may have differentiated—that is, melted and separated into layers of different density and composition, as did the Terrestrial planets. The fragmentation of such a planetesimal could produce different types of meteorites.

and produce different kinds of meteorites (Figure 12-4). In contrast, chondrites are probably fragments of smaller bodies that never melted, and carbonaceous chondrites may be from unaltered bodies that formed especially far from the sun.

SCIENTIFIC ARGUMENT

How can you say that meteors come from comets, but meteorites come from asteroids?

To begin with, remember the distinction between meteors and meteorites. A meteor is the streak of light seen in the sky when a particle from space is heated by friction with Earth's atmosphere. A meteorite is a piece of space material that actually reaches the ground.

The distinction between comet and asteroid sources must take into account two very strong effects that prevent you from finding meteorites that originated in comets. First, evidence that you will learn more about later in this chapter indicates that cometary material is physically weak, so comet particles vaporize in Earth's atmosphere easily. Very few ever reach the ground. Second, even if a comet particle reached the ground, it would be so fragile that it would weather away rapidly, and you would be unlikely to find it before it disappeared. Asteroidal particles, however, are made from rock and metal and are stronger. They are more likely to survive their plunge through the atmosphere and, afterwards, more likely to survive erosion on the ground. Every known meteorite is from the asteroids—not a single meteorite is known to be cometary. In contrast, meteor tracks show that most meteors you see come from comets, and very few are from the asteroid belt.

Now build a new argument. **What evidence suggests that meteorites were once part of larger bodies broken up by impacts?**

12-2 Asteroids

ASTEROIDS ARE DISTANT OBJECTS too small to study in detail with Earth-based telescopes. Astronomers nevertheless have learned a surprising amount about those little worlds using spacecraft and space telescopes.

Properties of Asteroids

Evidence from meteorites shows that the asteroids are the last remains of the rocky planetesimals that built the Terrestrial planets 4.6 billion years ago.

Study **Observations of Asteroids** on pages 244–245, noticing four important points and three new terms:

1. Most asteroids are irregular in shape and battered by impact cratering. Many asteroids seem to be rubble piles of broken fragments.
2. Some asteroids are double objects or have small moons in orbit around them. This is further evidence that asteroids have suffered collisions.

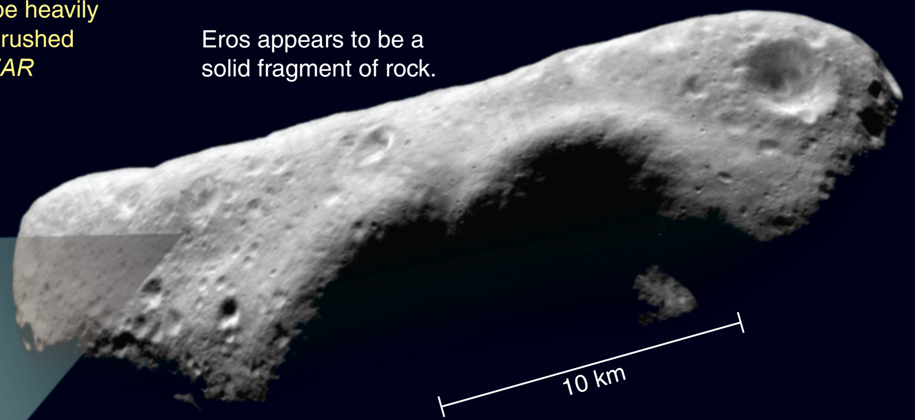
Observations of Asteroids

1 Seen from Earth, asteroids look like faint points of light moving in front of distant stars. Not many years ago they were known mostly for drifting slowly through the field of view and spoiling long time exposures. Some astronomers referred to them as “the vermin of the sky.” Spacecraft have now visited asteroids, and the images radioed back to Earth show that the asteroids are mostly small, gray, irregular worlds heavily cratered by impacts.

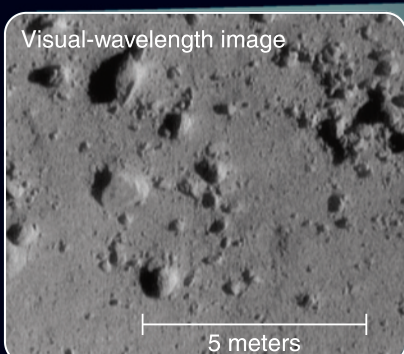
The *Near Earth Asteroid Rendezvous (NEAR)* spacecraft visited the asteroid Eros in 2000 and found it to be heavily cratered by collisions and covered by a layer of crushed rock ranging from dust to large boulders. The *NEAR* spacecraft eventually landed on Eros.

Visual-wavelength image

Eros appears to be a solid fragment of rock.



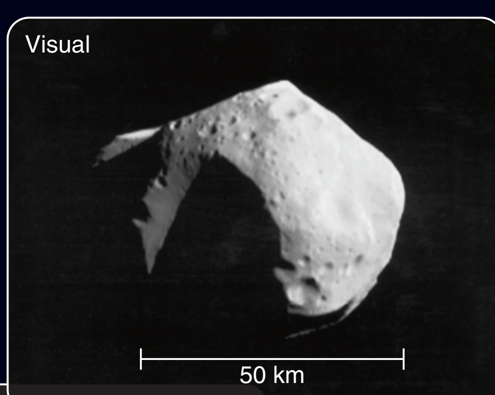
NASA/JPL/JHUAPL



NASA/JPL/JHUAPL

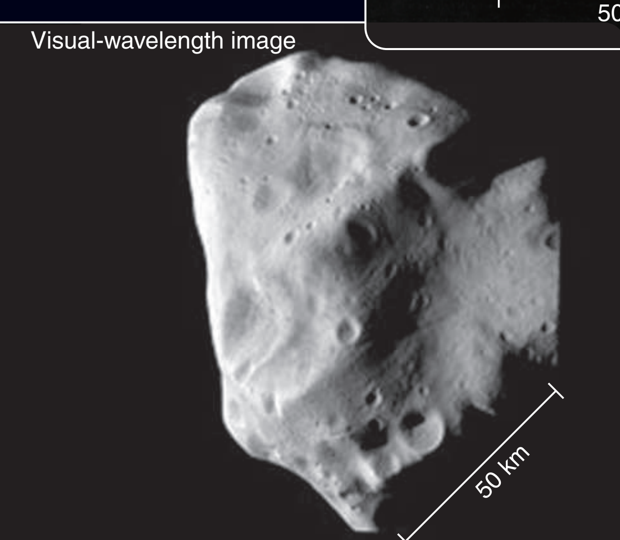
Most asteroids are too small for their gravity to pull them into a spherical shape. Impacts break them into irregularly shaped fragments.

NASA



The surface of Mathilde is very dark rock.

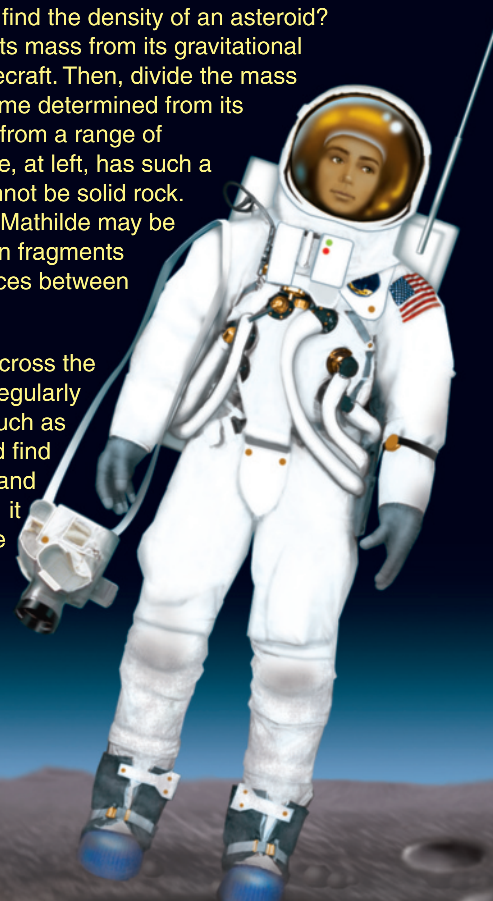
1a How would you find the density of an asteroid? First, measure its mass from its gravitational pull on passing spacecraft. Then, divide the mass by the asteroid’s volume determined from its size in images made from a range of perspecitves. Mathilde, at left, has such a low density that it cannot be solid rock. Like many asteroids, Mathilde may be a rubble pile of broken fragments with large empty spaces between fragments.



ESA/MPS/UPD/LAM/IAA/RSSD/INTA/UPM/DASP/IDA

Asteroid Lutetia, photographed by the *Rosetta* spacecraft at closest approach, is a large asteroid with an unusual spectrum and high density that lead astronomers to speculate it originated in the Terrestrial planet zone and was somehow tossed into the main asteroid belt.

If you walked across the surface of an irregularly shaped asteroid such as Eros, you would find gravity very weak; and in many places, it would not be perpendicular to the surface.

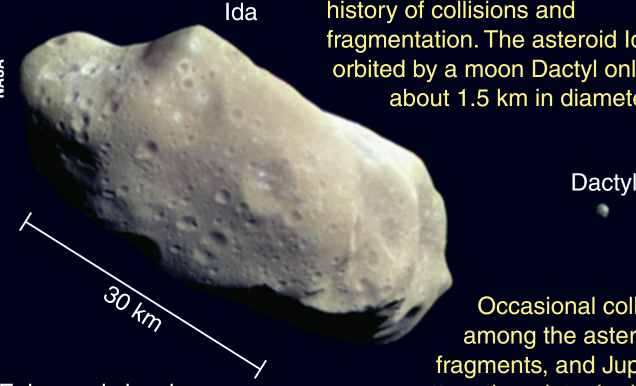


Line art on this page © Cengage Learning 2014


Radar image

NASA

2 Asteroids that pass near Earth can be imaged by radar. The asteroid Toutatis is revealed to be a double object—two objects orbiting close to each other or actually in contact.



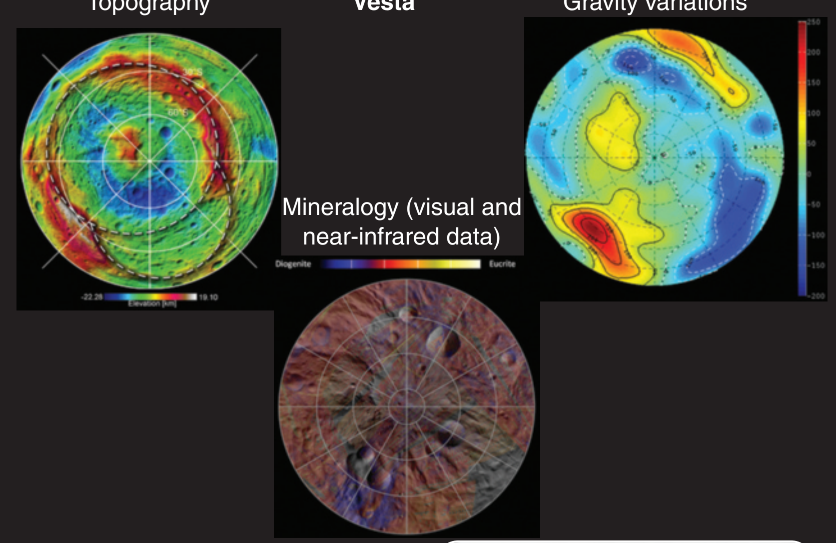
NASA

Enhanced visual + infrared

Double asteroids are more common than was once thought, reflecting a history of collisions and fragmentation. The asteroid Ida is orbited by a moon Dactyl only about 1.5 km in diameter.

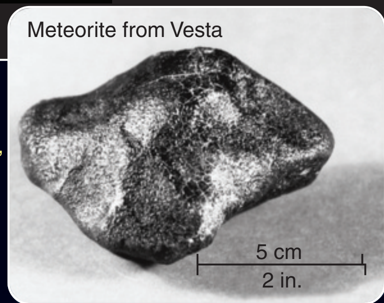
Occasional collisions among the asteroids release fragments, and Jupiter's gravity scatters them into the inner solar system as a continuous supply of meteorites.

NASA/JPL-Caltech/UCLA/INAF/MPS/DLR/IDA



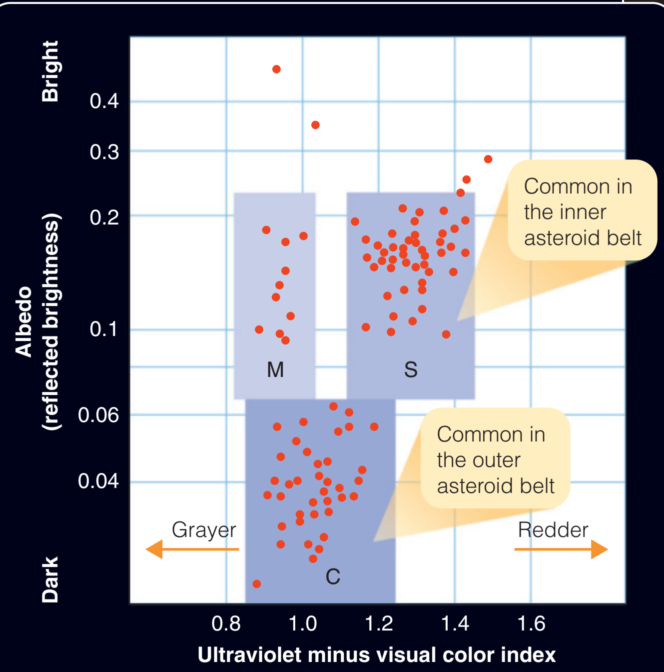
3 The upper left and upper right images, based on data acquired by the orbiting *Dawn* spacecraft, respectively show the topography of Vesta's southern hemisphere and a map of Vesta's gravity variations adjusted for Vesta's shape, revealing variations in density. This is the same portion of Vesta shown in the chapter opening image on page 238. The central peak of the large Rheasilvia basin, which appears as the yellow area just above and to the left of center, has a small positive gravity anomaly, indicating that the material there is denser, perhaps originally coming from deeper within the body.

The lower-center image shows the mineral distribution in Vesta's southern hemisphere crust from *Dawn's* visible and infrared mapping spectrometer, which captures different wavelengths of reflected and emitted radiation. The areas in purple have a higher proportion of diogenite minerals, and yellow areas have a higher proportion of eucrite minerals. Geologists interpret the patterns in this image to suggest that Vesta likely melted all the way through early in its history.



Lab photograph, R. Kempton, New England Meteoritical/NASA

3a A class of meteorites that are spectroscopically identical to Vesta are thought to be fragments of this asteroid, perhaps blasted into space by the collision that created the big southern basin. Those meteorites appear to be solidified basalt lava. The spectrum of Vesta and the composition of its associated meteorites provide evidence that this asteroid once had geological activity.



4 Although asteroids would look gray to your eyes, they can be classified according to their albedos (reflected brightness) and spectroscopic colors. As shown at left, **S-types** are brighter and tend to be reddish. They are the most common kind of asteroid and appear to be the source of the most common chondrite meteorites.

M-type asteroids are not too dark but are also not very red. They may be mostly iron-nickel alloys.

C-type asteroids are as dark as lumps of sooty coal and appear to be carbonaceous.

3. A few asteroids show signs of geological activity that probably happened on their surfaces when those asteroids were young.
4. Asteroids can be classified by their albedos, colors, and spectra to reveal clues to their compositions. This also allows them to be compared to meteorites in labs on Earth. *C-type, S-type, and M-type* asteroids are the main classes discovered by this method.

The Asteroid Belt

The first asteroid was discovered in 1801 by the Sicilian monk Giuseppe Piazzi and later named Ceres. Astronomers were excited by Piazzi's discovery because Ceres orbits 2.8 AU from the sun in a wide gap between Mars and Jupiter where it was thought a planet "ought" to be, but Ceres has less than ¼ the diameter of Earth's moon, much smaller than any of the planets. Three even smaller objects—Pallas, Juno, and Vesta—were discovered within a few years, all orbiting between Mars and Jupiter, so astronomers decided that Ceres and the other asteroids should not be considered true planets. You learned in Chapter 11 that Ceres, the largest object in the asteroid belt, has now been reclassified as a dwarf planet because it has enough gravitational strength to squeeze itself into a spherical shape but not enough to have swept up or cleared away the rest of the asteroids.

Today over 300,000 asteroids have well-charted orbits. Only three are larger than 400 km (250 mi) in diameter; most are much smaller. There are probably a million or more that are larger than 1 km (0.6 mi) in diameter. Astronomers are sure that all the large asteroids in the asteroid belt have been discovered but are also sure that many small asteroids remain undiscovered.

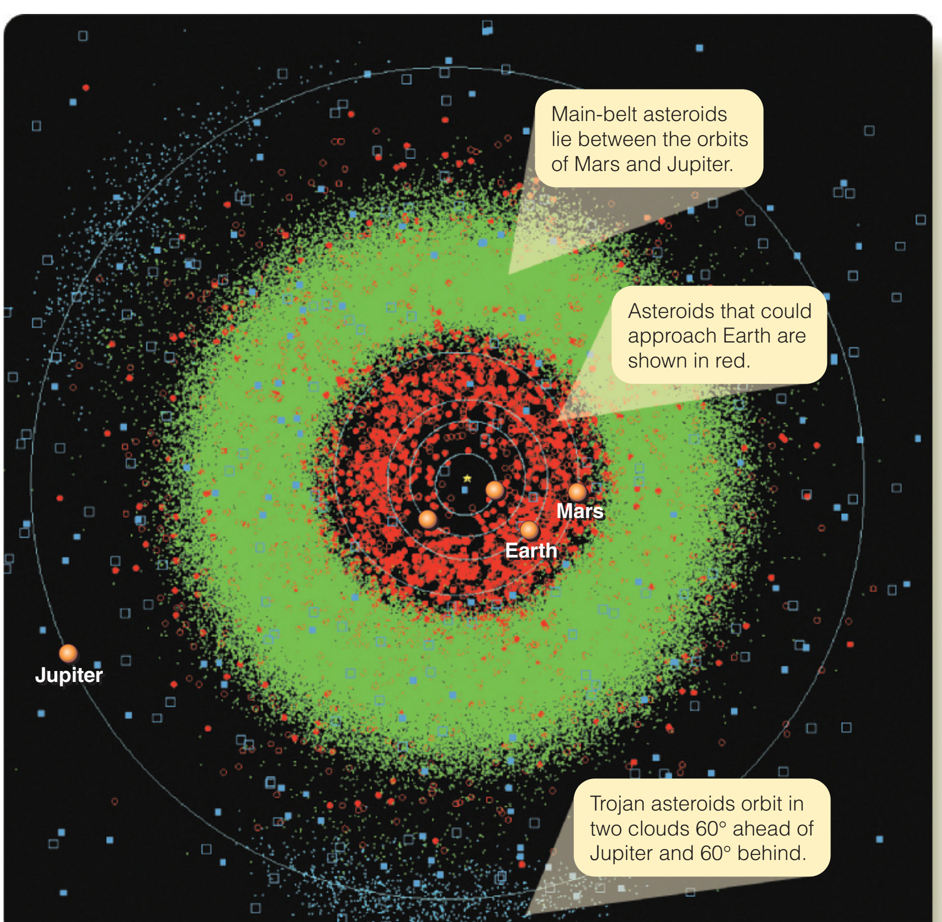
Movies and TV have created a **Common Misconception** that flying through an asteroid belt is a hair-raising plunge requiring constant dodging left and right. The asteroid belt between Mars and Jupiter is actually mostly empty space. In fact, if you were standing on an asteroid, it would be many months or years between sightings of other asteroids.

Asteroids Outside the Main Belt

You don't have to go all the way to the asteroid belt if you want to visit an asteroid; some follow orbits that cross the orbits of the Terrestrial planets and come near Earth. Others wander far away, among the Jovian worlds. There are also asteroids that share orbits with the planets (■ Figure 12-5).

Several research teams are now intent on identifying **Near-Earth Objects (NEOs)** which could potentially hit Earth. The most dangerous type of NEO, those with orbits that cross Earth's orbit, are called **Apollo objects** after the first such asteroid discovered. The combined searches are expected to be able to locate at least 90 percent of the Apollos and other NEOs larger than 1 km in diameter by 2013.

In March 2009, NEO searchers spotted a small asteroid (2 to 3 meters in diameter, about the size of a small truck) on a collision course with Earth. Astronomers were able to observe it in space before impact and discovered that its color and albedo matched a type of asteroid found mostly in the outer belt. The asteroid entered Earth's atmosphere over the deserts of northern Sudan and exploded. Scientists from the U.S. and Sudan searched for pieces of the object (■ Figure 12-6).



© Cengage Learning 2014; Images: Minor Planet Center/Harvard University

■ **Figure 12-5**

This diagram plots the position of known asteroids between the sun and the orbit of Jupiter on a specific day. Most asteroids are in the main belt. Squares, filled or empty, show the location of known comets. Although asteroids and comets are small bodies and lie far apart, there are a great many of them in the inner solar system.

P.Jenniskens/SETI Institute

■ **Figure 12-6**

Search team leader Muawia Shaddad (left) of the Faculty of Physics and Sciences at the University of Khartoum, and meteor astronomer Peter Jenniskens (right) of the Carl Sagan Center at the SETI Institute pose with one of the first fragments found of the small asteroid that exploded over the Sudanese desert in March 2009.

They found about 4 kg (9 lb) of fragments corresponding to a rare type of stony meteorite. For the first time, a definite connection could be made between an asteroid observed in space and meteorites with properties measured in Earth laboratories. In 2010, the Japanese probe *Hayabusa* returned a capsule to Earth containing a few tiny grains from the surface of the S-type asteroid Itokawa, which were found to be chemically identical to one of the most common types of ordinary chondrite meteorites.

There are other nonbelt asteroids beyond the main belt. Gravitational effects of Jupiter and the sun combine to trap two groups of asteroids in the Lagrange points that are 60° ahead of and 60° behind the planet in its orbit (Figure 12-5). These are called **Trojan asteroids** because individual objects have been named after heroes of the Trojan War. Astronomers have also found a few objects in the Lagrange points of the orbits of Mars and Neptune.

The object Chiron, found in 1977, is about 170 km (110 mi) in diameter. Its orbit carries it from the orbit of Uranus to just inside the orbit of Saturn. Objects like Chiron with orbits between, or crossing, orbits of the Jovian planets are called **centaurs**. Although it was first classified as an asteroid, Chiron later surprised astronomers by suddenly releasing jets of vapor and dust. You will learn in the next section that this more resembles the characteristics of comets than asteroids. Centaurs show that the distinction between asteroids and comets is not clear-cut.

As technology allows astronomers to detect smaller and more distant objects, they are learning that our solar system contains large numbers of these small bodies. The challenge is to explain their origin.

Origin and History of the Asteroids

An old hypothesis proposed that asteroids are the remains of a planet that exploded. Planet-shattering death rays may make for exciting science-fiction movies, but in reality planets do not explode. The gravitational field of a planet holds the mass together so tightly that completely disrupting the planet would take tremendous energy. In addition, the present-day total mass of the asteroids is only about 1/20 the mass of Earth's moon, hardly enough to be the remains of a planet.

Astronomers have evidence that the asteroids are the remains of material lying 2 to 4 AU from the sun that was unable to form a planet because of the gravitational influence of Jupiter, the next planet outward. Over the 4.6-billion-year history of the solar system, most of the objects originally in the asteroid belt have collided, fragmented, and been covered with craters. Some asteroids were perturbed by the gravity of Jupiter and other planets into orbits that collided with planets or the sun, caused them to be captured as planetary satellites, or ejected them from the solar system. The present-day asteroids are understood to be a very minor remnant of the original mass in that zone.

Even though most of the planetesimals originally in the main belt have been lost or destroyed, the objects left behind carry clues to their origin in their albedos and spectroscopic colors (see p. 245). For example, C-type asteroids have albedos less than 0.06 and would look very dark to your eyes. They are probably made of carbon-rich material similar to that in carbonaceous chondrite meteorites. C-type asteroids are more common in the outer asteroid belt. It is cooler there, and the condensation sequence (look back to Chapter 8) predicts that carbonaceous material would form there more easily than in the inner belt.

You learned that the third largest asteroid, Vesta, has been geologically active, with evidence of solidified lava on its surface (p. 245). But, not all large asteroids have been active. Ceres, 900 km in diameter, is almost twice as big as Vesta, but it shows no spectroscopic sign of past activity and evidently has an ice-rich mantle. The puzzling differences between those two large asteroids will be investigated by the *Dawn* spacecraft that orbited Vesta from 2011 to 2012 and then departed and is expected to arrive and begin orbiting Ceres in 2015.

Although there are still mysteries to solve, you can understand the story of the asteroids. They are fragments of planetesimals, some of which differentiated, developed molten metal cores, in a few cases even had lava flows on their surfaces, and then cooled slowly. The largest asteroids astronomers see today may be nearly unbroken examples of original planetesimals, but the smaller asteroids are fragments produced by 4.6 billion years of collisions.

SCIENTIFIC ARGUMENT

What is the evidence that asteroids have been fragmented?

First, your argument might note that the solar nebula theory predicts that planetesimals collided and either stuck together or fragmented. This is suggestive, but it is not evidence. A theory can never be used as evidence to support some other theory or hypothesis. Evidence means observations or the results of experiments, so your argument must cite observations. The spacecraft photographs of asteroids show irregularly shaped little worlds heavily scarred by impact craters. Further evidence indicates some asteroids may be pairs of bodies split apart but still in contact, and images of asteroid Ida reveal a small satellite, Dactyl. Other asteroids with moons have been found. These double asteroids and asteroids with moons probably reveal the results of fragmenting collisions between asteroids. Furthermore, meteorites appear to have come from the asteroid belt astronomically recently, so fragmentation must be a continuing process there.

Now build an argument to combine what you know of meteorites with your experience with asteroids. **What evidence could you cite to show what the first planetesimals were like?**

12-3 Comets

Few sights in astronomy are more beautiful than a bright comet hanging in the night sky (■ Figure 12-7). It is a **Common Misconception** that comets whiz rapidly across the sky like meteors. Actually, a comet's motion is hardly apparent; night by night its position shifts slightly against the background stars, and it may remain visible for months. While everyone can enjoy the beauty of comets, astronomers study them because they are messengers from the past carrying cargos of information about the origin of our solar system.

Properties of Comets

As always, you should begin your study of a new kind of object by summarizing its observational properties. What do comets look like, and how do they behave?

Study **Observations of Comets** on pages 250–251 and notice three important properties of comets plus three new terms:

1. Comets have two kinds of tails, shaped by the solar wind and solar radiation. Gas and dust released by a comet's icy nucleus produces a head, or *coma,* and are then blown outward, away from the sun. The gas produces a *gas tail,* and the dust produces a separate *dust tail.*
2. Comet dust produces not only one of the two types of comet tails but also spreads throughout the solar system. Some of those comet dust particles later encounter the Earth and are seen as meteors.
3. Evidence shows that some comet nuclei are fragile and can break into pieces.

Astronomers can put these and other observations together to study the structure of comet nuclei.

The Geology of Comet Nuclei

The nuclei of comets are quite small and cannot be studied in detail using Earth-based telescopes. Nevertheless, when a comet nucleus approaches the sun, it emits material that forms into a coma (head) and tail that can be millions of kilometers in size and is easily observed (look back to Chapter 8, page 148).

Spectra of comet comae (plural of *coma*) and tails indicate the nuclei must contain ices of water and other volatile compounds such as carbon dioxide, carbon monoxide, methane, ammonia, and so on. These are the kinds of compounds that would have condensed in cold regions of the solar nebula. This convinces astronomers that comets are ancient samples of the gases and dust from which the outer planets formed. As the ices absorb energy from sunlight they sublime—change from a solid directly into a gas. The gases break down and also combine chemically, producing other substances found in comet spectra. For example, vast clouds of hydrogen gas observed around the heads of comets are understood to derive from the breakup of ice molecules.

Five spacecraft flew past the nucleus of Comet Halley when it visited the inner solar system in 1985 and 1986. Other spacecraft flew past the nuclei of Comet Borrelly in 2001, Comet Wild 2 (pronounced *vildt-two*) in 2004, Comet Tempel 1 in 2005, and Comet Hartley 2 in 2010. Images show that all these comet nuclei are 1 to 10 km across, similar in size to many asteroids, and also irregular in shape (■ Figure 12-8). In general, these nuclei are darker than a lump of coal, which suggests composition similar to carbon-rich carbonaceous chondrite meteorites described earlier in this chapter.

The mass and density of comet nuclei can be calculated from their gravitational influence on passing spacecraft. Comet nuclei appear to have densities between 0.1 and 0.25 g/cm^3, much less than the density of ice. Also, as you will learn later in this chapter, comets subjected to tidal stresses from Jupiter or the sun come apart very easily. Comet nuclei have been described as dirty snowballs or icy mudballs, but that seems to be incorrect; their shapes, low densities, and lack of material strength suggest that comets are not solid objects. The evidence leads astronomers to conclude that most comet nuclei must be fluffy mixtures of ices and dust with significant amounts of empty space. On the other hand, images of the nucleus of Comet Wild 2 revealed cliffs, pinnacles, and other features that show the material has enough strength to stand against the weak gravity of the comet.

The nuclei of comets appear to have a crust of rocky dust left behind as the ices vaporize. Breaks in that crust can expose ices

© John White

■ **Figure 12-7**

Comet McNaught swept through the inner solar system in 2007 and was a dramatic sight in the southern sky. Seen here from Australia, the comet was on its way back into deep space after making its closest approach to the sun ten days earlier. Comet McNaught began this passage with a period of about 300,000 years, but gravitational perturbations by the planets changed its orbit shape from elliptical to hyperbolic, so it will never return but instead is leaving the solar system to journey forever in interstellar space.

to sunlight, and vents can occur in those regions. It also seems that some comets have large pockets of volatiles buried below the crust. When one of those pockets is exposed and begins to vaporize, the comet can suffer a dramatic outburst, as Comet Holmes did in 2007 (■ Figure 12-9).

Astronomers have devised ways to study comet material more directly. The *Stardust* spacecraft passed through the tail of Comet Wild 2 in 2004, collected dust particles that had been ejected from the comet's nucleus, and returned the samples in a sealed capsule to Earth in 2006 for analysis. In 2005, the *Deep Impact* spacecraft released an instrumented impactor probe into the path of comet Tempel 1. As planned, the nucleus of the comet ran into the impactor at almost 10 km/s (22,000 mph). The probe broke through the crust of the nucleus and blasted vapor and dust out into space where the *Deep Impact* "mother ship," as well as the *Spitzer* and *Hubble* space telescopes and observatories on Earth, could analyze it (see page 251). Those missions produced the surprising discovery that some comet dust is crystalline and must have formed originally in very warm environments close to the sun, but then was incorporated somehow into comet nuclei in the cold outer solar system.

Origin and History of Comets

Family relationships among the comets can give you clues to their origin. Most comets have long, elliptical orbits with periods greater than 200 years and are known as long-period comets. The long-period comet orbits are randomly inclined to the plane of the solar system, so those comets approach the inner solar system from all directions. Long-period comets revolve around the sun in about equal numbers in prograde orbits (the same direction in which the planets move) and retrograde orbits.

Observations of Comets

1 A **gas tail** is produced by ionized gas carried away from the nucleus by the solar wind. The spectrum of a gas tail is an emission spectrum. The atoms are ionized by the ultraviolet light in sunlight. The wisps and kinks in gas tails are produced by the magnetic field embedded in the solar wind.

Spectra of gas tails reveal atoms and ions such as H_2O, CO_2, CO, H, OH, O, S, C, and so on. These are released by the vaporizing ices or produced by the breakdown of those molecules. Some gases, such as hydrogen cyanide (HCN), must be formed by chemical reactions.

Gas tail

Dust tail

1a A **dust tail** is produced by dust that was contained in the vaporizing ices of the nucleus. The dust is pushed gently outward by the pressure of sunlight, and it reflects an absorption spectrum, the spectrum of sunlight. The dust is not affected by the magnetic field of the solar wind, so dust tails are more uniform than gas tails. Dust tails are often curved because the dust particles follow their individual orbits around the sun once they leave the nucleus. Because of the forces acting on them, both gas and dust tails extend away from the sun.

Nucleus

1b The nucleus of a comet (too small to be visible here) is a lump of fragile, porous material containing ices of water, carbon dioxide, ammonia, and so on. Comet nuclei can be 1 to 100 km in diameter.

Coma

The **coma** of a comet is the cloud of gas and dust that surrounds the nucleus. It can be over 1,000,000 km in diameter, bigger than the sun.

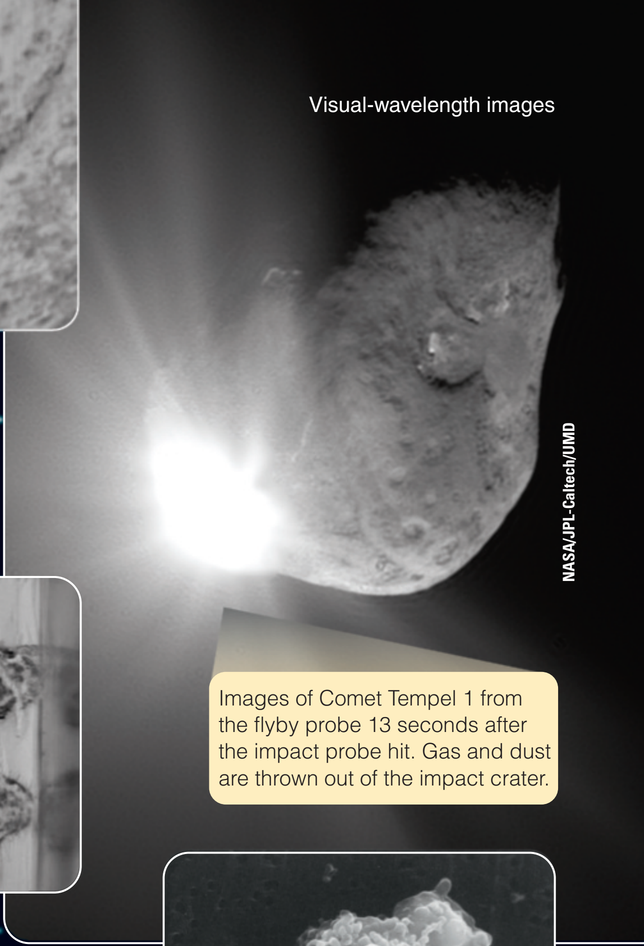
1c Comet Mrkos in 1957 shows how the gas tail can change from night to night due to changes in the magnetic field in the solar wind.

Visual-wavelength images

2 As the ices in a comet nucleus vaporize, they release dust particles that not only form the dust tail but also spread throughout the solar system.

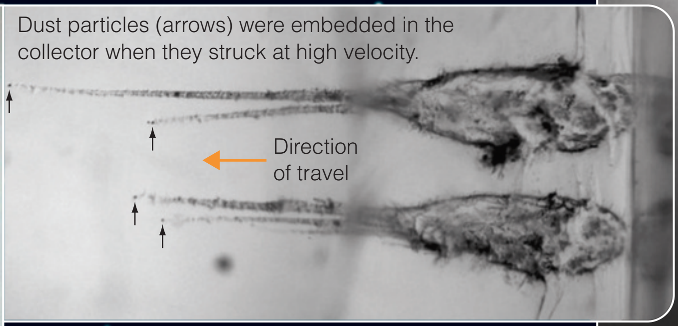
The *Deep Impact* spacecraft released an instrumented probe into the path of Comet Tempel 1. When the comet slammed into the probe at 10 km/s as shown at right, huge amounts of gas and dust were released. From the results, scientists conclude that the nucleus of the comet is rich in dust finer than the particles of talcum powder. The nucleus is marked by craters, but it is not solid rock. It is about the density of fresh-fallen snow.

Only seconds before impact, craters are visible on the dark surface.



Images of Comet Tempel 1 from the flyby probe 13 seconds after the impact probe hit. Gas and dust are thrown out of the impact crater.

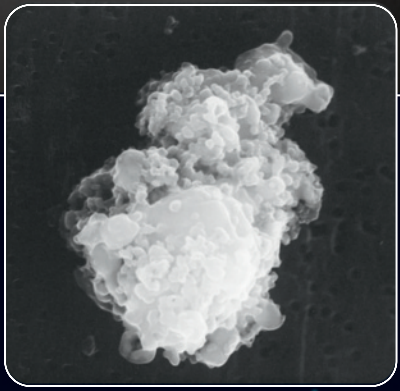
A microscopic mineral crystal from Comet Wild 2



Dust particles (arrows) were embedded in the collector when they struck at high velocity.

2a The *Stardust* spacecraft flew past the nucleus of Comet Wild 2 and collected dust particles (as shown above) in an exposed target that was later parachuted back to Earth. The dust particles hit the collector at high velocity and became embedded, but they can be extracted for study.

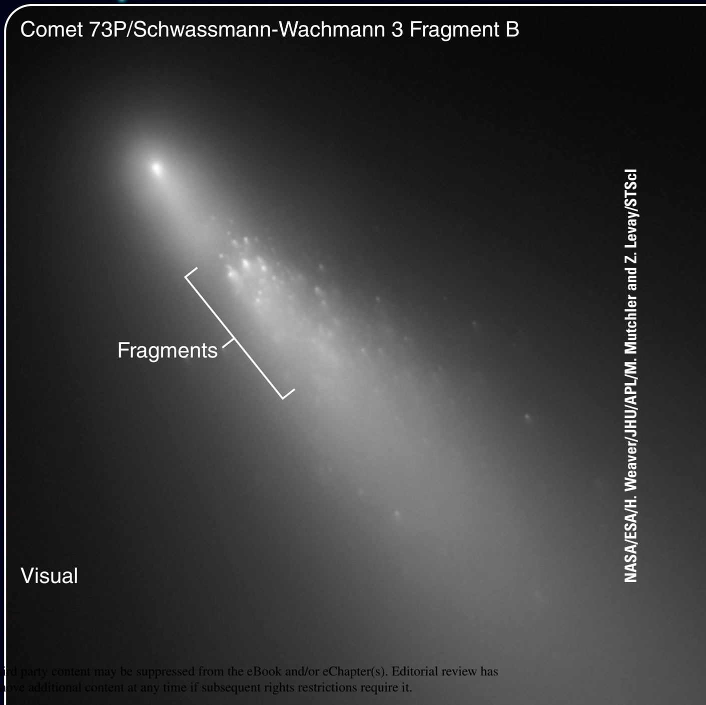
Some of the collected dust is made of high-temperature minerals that could only have formed near the sun. This suggests that material from the inner solar nebula was mixed outward and became part of the forming comets in the outer solar system. Other minerals found include olivine, a very common mineral on Earth but not one that scientists expected to find in a comet.



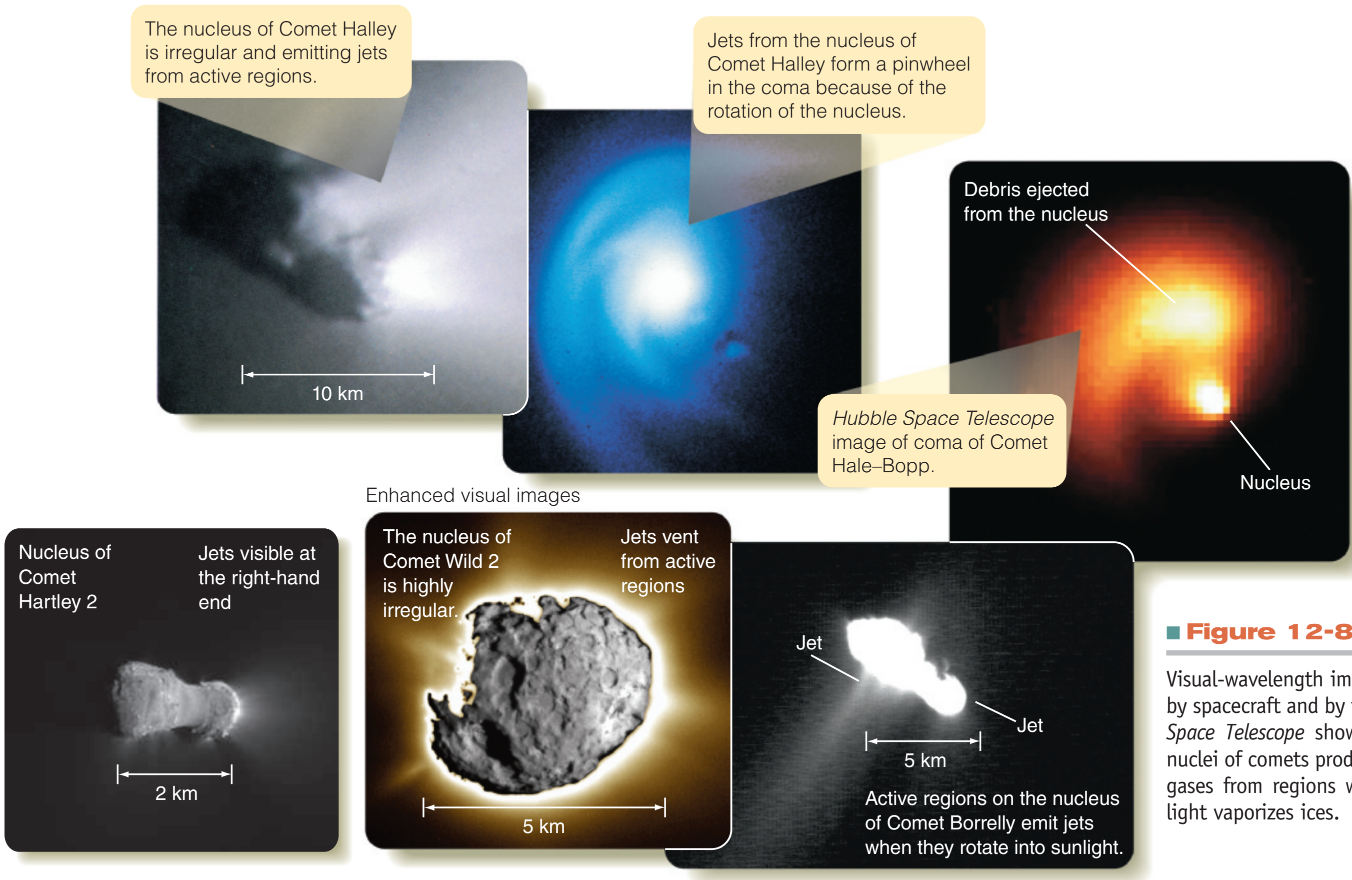
This dust particle was collected by spacecraft above Earth's atmosphere. It is almost certainly from a comet.

3 The nuclei of comets are not strong and can break up. In 2006, Comet Schwassmann-Wachmann 3 broke into a number of fragments that themselves fragmented. Fragment B is shown at the right breaking into smaller pieces. The gas and dust released by the breakup made the comet fragments bright in the night sky, and some were visible with binoculars. As its ices vaporize and its dust spreads, the comet may totally disintegrate and leave nothing but a stream of debris along its previous orbit.

Comets most often break up as they pass close to the sun or close to a massive planet like Jupiter. Comet LINEAR broke up in 2000 as it passed by the sun. Comet Shoemaker-Levy 9 that hit Jupiter in 1994 was first ripped to pieces by tidal stresses from Jupiter's gravity. Comets can also fragment far from planets, perhaps because of the collapse of cavities within the icy nucleus.



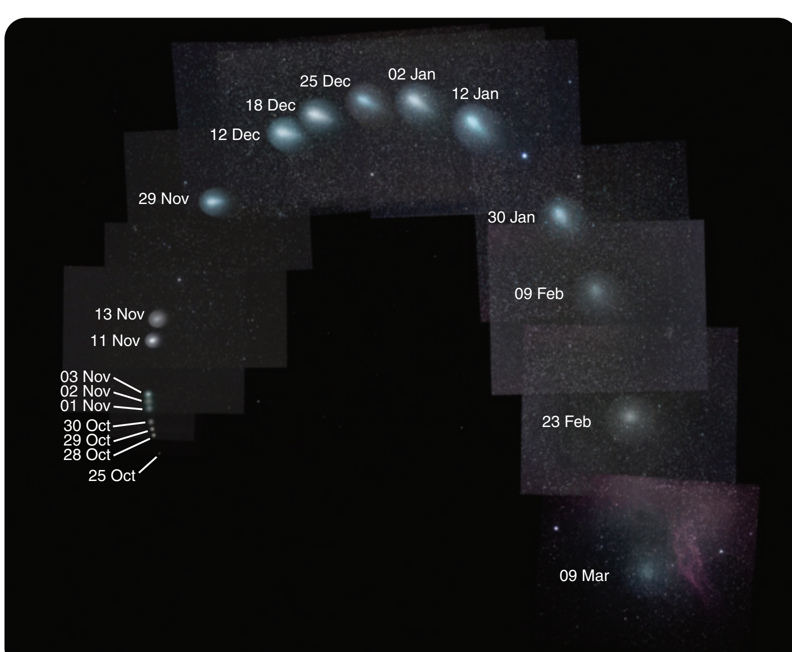
Line art on this page © Cengage Learning 2014



© Cengage Learning 2014; Images: Halley nucleous: © 1996 Max Planck Institute; Halley coma: Steven Larson; Comets Borelly, Hale-Bopp, Wild 2, and Hartley 2: NASA

Figure 12-8

Visual-wavelength images made by spacecraft and by the *Hubble Space Telescope* show how the nuclei of comets produce jets of gases from regions where sunlight vaporizes ices.



Photos by John F. Pane

Figure 12-9

Composite of 19 snapshots of Comet Holmes, showing its changing brightness and position spanning the period from October 2007 until March 2008. During its outburst in late October 2007, the comet brightened by a factor of about 500,000 as a large pocket of volatile material exploded through its crust and spread into space.

In contrast, about 100 or so of the 600 well-studied comets have orbits with periods less than 200 years. These short-period comets usually follow orbits that lie within 30° of the plane of the solar system, and most revolve around the sun prograde. Comet Halley, with a period of 76 years, is an unusual short-period comet with a retrograde orbit.

A comet cannot survive long in an orbit that brings it into the inner solar system. The heat of the sun vaporizes millions of tons of ices during each passage and reduces comets to inactive bodies of rock and dust; such comets can last at most 100 to 1000 orbits around the sun. Astronomers calculate that even before a comet completely vaporizes from solar heating, it can't survive more than about half a million years crossing the orbits of the planets, especially Jupiter, without having its path rerouted into the sun or out of the solar system or colliding with one of the planets. Therefore, comets visible in our skies now can't have survived in their present orbits for 4.6 billion years since the formation of the solar system, and that means there must be a continuous supply of new comets. Where do they come from?

Comets from the Oort Cloud and the Kuiper Belt

In the 1950s, astronomer Jan Oort proposed that the long-period comets are objects that fall inward from what has become known as the **Oort cloud,** a spherical cloud of icy bodies that extends from about 10,000 to 100,000 AU from the sun. Astronomers hypothesize that the Oort cloud objects were originally planetesimals formed in the outer solar system that were tossed to their current distant locations by the gravitational influence of the growing Jovian planets. The Oort cloud is estimated to contain several trillion (10^{12}) icy bodies. Far from the sun, they are very cold, lack comae and tails, and are invisible from Earth.

Astronomers calculate that some Oort cloud objects can be perturbed by the gravitational influence of passing stars so that they fall into the inner solar system. There, the heat of the sun warms their ices and transforms them into comets. The fact that long-period comets are observed to fall inward from all directions is explained by their Oort cloud reservoir being spherically distributed around the sun and inner solar system. A few short-period comets, including Comet Halley, also seem to have originated as long-period comets from the Oort cloud but then had their orbits altered by a close encounter with Jupiter. However, that process can't explain all of the short-period comets.

The Kuiper belt, which you encountered in Chapter 8 when you studied the origin of the solar system, is another source for icy bodies passing through the inner solar system. Kuiper belt objects (KBOs) orbit in the same direction as the planets and in the plane of the solar system. Mathematical models show that it is possible for a KBO perturbed inward by the influence of the giant planets to move into an orbit resembling those of the short-period comets.

Comets vary in brightness and orbit. Nevertheless, there are two basic types of comets in our solar system. Some originate in the Oort cloud far from the sun. Others come from the Kuiper belt just beyond Neptune. They all share one characteristic—they are ancient icy bodies that were born when the solar system was young.

SCIENTIFIC ARGUMENT

How do comets help explain the formation of the planets?
This argument must refer to the solar nebula hypothesis. The planetesimals that formed in the inner solar nebula were warm and could not incorporate much ice. The asteroids are understood to be the last remains of such rocky bodies. On the other hand, planetesimals in the outer solar system contained large amounts of ices. Many of them were destroyed by being accreted together to make the Jovian planets, but some survived intact. The icy bodies of the Oort cloud and the Kuiper belt may be the solar system's last surviving icy planetesimals. When those icy objects have their orbits perturbed by the gravity of the planets or passing stars, some are redirected into the inner solar system where you see them as comets. The gases released by comets indicate that they are rich in volatile materials such as water and carbon dioxide. These are the ices you would expect to find in the icy planetesimals. Comets also contain dust with rocklike chemical composition, and the planetesimals must have included large amounts of such dust frozen into the ices when they formed. Thus, the nuclei of comets seem to be frozen samples of the original outer solar nebula.

Nearly all of the mass of a comet is in the nucleus, but the light you see comes from the coma and the tail. Build a new argument to discuss observations. **What do spectra of comets tell you about the process that converts dirty ice into a comet?**

12-4 Asteroid and Comet Impacts

METEORITE IMPACTS ON HOMES of the type described at the start of this chapter are not common. Most meteors are small particles ranging from a few centimeters down to microscopic dust. Astronomers estimate that Earth gains about 40,000 tons of mass per year from meteorites of all sizes. (That may seem to you like a lot, but it is less than a hundred-thousandth of a trillionth of Earth's total mass.) Statistical calculations indicate that a meteorite large enough to cause some damage, like the one that hit Mrs. Hodges and her house in 1954, strikes a building somewhere in the world about once every 16 months. Declassified data from military satellites show that Earth is hit about once a week by meter-size asteroids. What happens when even larger solar system objects collide with Earth?

■ **Figure 12-10**

(a) The Barringer Meteorite Crater (near Flagstaff, Arizona) is nearly a mile in diameter and was formed about 50,000 years ago by the impact of an iron meteorite estimated to have been roughly 50 meters in diameter. Notice the raised and deformed rock layers all around the crater. The brick museum building visible on the far rim at right provides some idea of scale. (b) Like all large-impact features, the Barringer Meteorite Crater has a raised rim and scattered ejecta.

USGS

Recent Hits

Barringer Crater near Flagstaff, Arizona, is 1.2 km (¾ of a mile) in diameter and 200 m (650 ft) deep (■ Figure 12-10). Eugene Shoemaker proved in his 1963 doctoral thesis that the crater must be the result of an impact rather than a volcanic eruption because quartz crystals in and around it had been subjected to pressures much higher than can be produced by a volcano. Further studies showed that Barringer Crater was created approximately 50,000 years ago by a meteorite estimated to have been about 50 m (160 ft) in diameter, as large as a good-sized building, that hit at a speed of 11 km/s, releasing as much energy as a large thermonuclear bomb. An object of that size could be called either a large meteorite or a small asteroid. Debris at the site shows that the impactor was composed of iron.

On a summer morning in 1908, reindeer herders and homesteaders in central Siberia were startled to see a brilliant blue-white fireball brighter than the sun streak across the sky. Still descending, it exploded with a blinding flash and an intense pulse of heat. The blast was heard up to 1000 km away, and the resulting air pressure circled Earth twice. When members of a scientific expedition arrived at the site in 1927, they found that the blast had occurred above the Stony Tunguska River valley and had knocked down trees pointing away from the center of a region about 30 km (20 mi) in radius (■ Figure 12-11).

No crater has been found at the Tunguska site, so it seems that the explosion, estimated to have equaled 12 megatons (12 million tons) of TNT, occurred at least a few kilometers above the ground. A detailed analysis of all the available evidence suggests that the impactor's speed and direction resembled the orbits of Apollo objects. Astronomers produced computer models of objects entering Earth's atmosphere and concluded that the most likely candidate for the Tunguska object is a stony asteroid about 30 m in diameter, perhaps one-tenth the mass of the Barringer impactor. An object of this size and material strength would have fragmented and exploded at just about the right height to produce the observed blast. This conclusion is consistent with modern studies of the Tunguska area showing that thousands of tons of powdered material with a composition resembling carbonaceous chondrites are scattered in the soil.

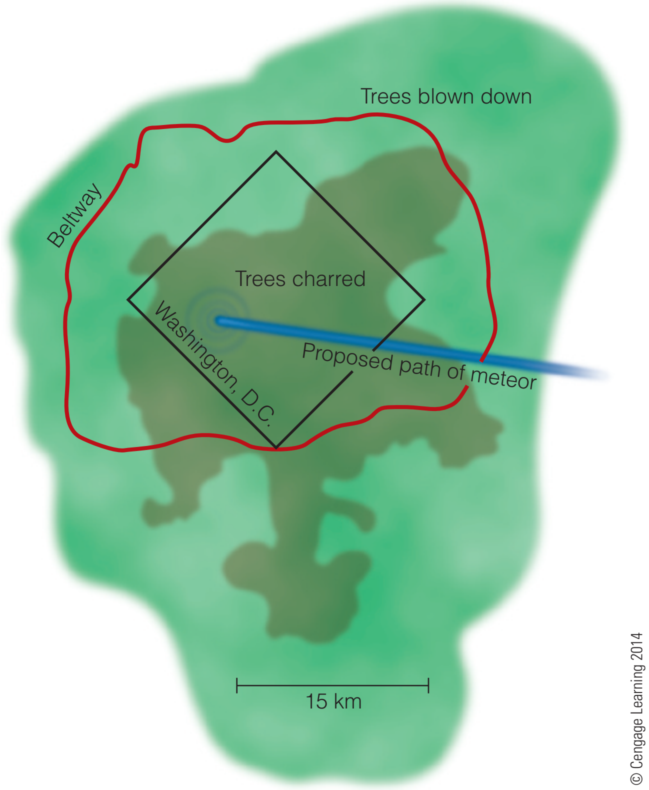
Planet-Shaking Events

There are some very big craters in the solar system, for example on the moon (look back to Chapter 9, page 179), that show what can happen when a full-sized asteroid or comet collides with a planet. Also, Earthlings watched in awe in 1994 as fragments from the nucleus of Comet Shoemaker-Levy 9 (abbreviated SL-9) slammed into Jupiter with a series of explosions equaling millions of megatons (that is, trillions of tons) of TNT (■ Figure 12-12). The impacts left scars larger than Earth visible on Jupiter's cloud tops for months afterward. In 2009 another object (it was not observed prior to impact) left a single blemish on Jupiter easily observed from Earth. Evidently impacts large enough to significantly affect a giant planet like Jupiter happen impressively often.

What would happen if an object the size of SL-9, or even larger, were to hit Earth? Sixty-five million years ago, at the end of the Cretaceous period, over 75 percent of the species on Earth,

M. A. Seeds

including the dinosaurs, became extinct. Scientists have found a thin layer of clay all over the world that was laid down at that time, and it is rich in the element iridium—common in meteorites but rare in Earth's crust. This suggests that an impact occurred that was large enough to have altered Earth's climate and caused the worldwide extinction.



© Cengage Learning 2014

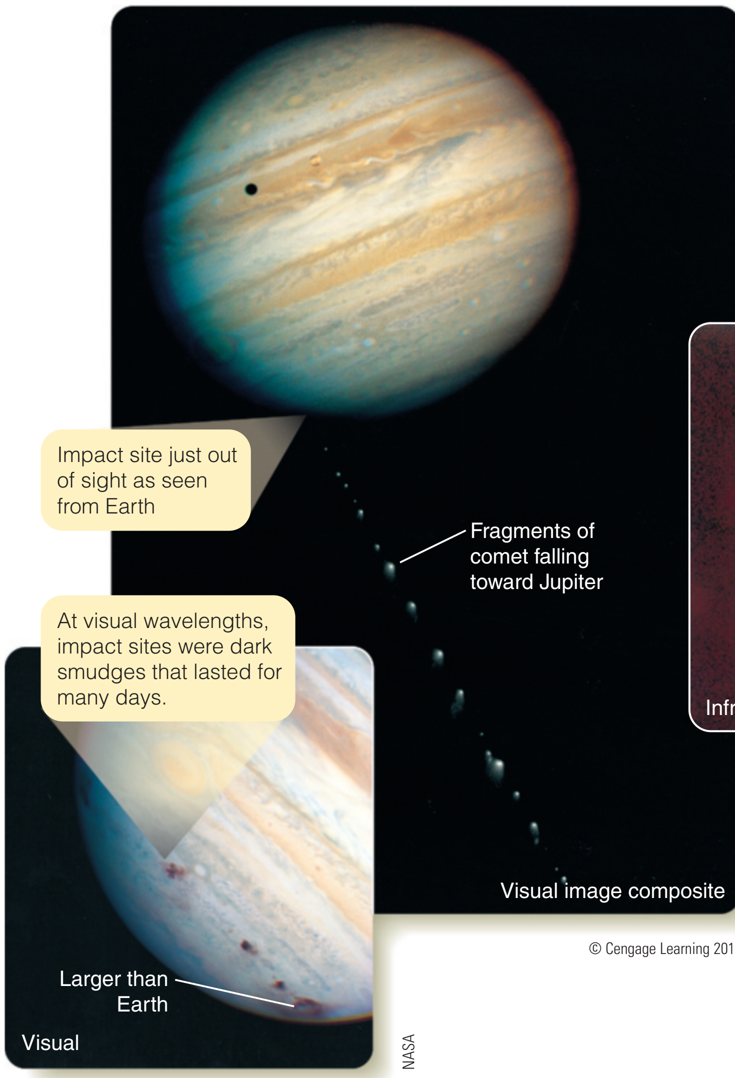
■ **Figure 12-11**

The 1908 Tunguska event in Siberia destroyed an area the size of a large city. Here the area of destruction is superimposed on a map of Washington, D.C., and its surrounding beltway. In the central area, trees were burned; in the outer area, trees were blown down in a pattern away from the path of the impactor.

Geologists have located a crater at least 180 km (110 mi) in diameter centered near the village of **Chicxulub** (pronounced *CHEEK-shoe-lube*) in the northern Yucatán region of Mexico (■ Figure 12-13). Although the crater is now completely covered by sediments, mineral samples show that it contains shocked quartz typical of impact sites and that it is the right age. The impact of an object 10 to 14 km in diameter formed the crater about 65 million years ago, just when the dinosaurs and many other species died out. Most scientists now conclude that this is the scar of the impact that ended the Cretaceous period.

There are a number of major extinctions in the fossil record, and at least some of these were probably caused by large impacts. Large asteroid impacts on Earth happen very rarely from a human perspective, but they happen often relative to geological and astronomical time scales. For example, astronomers estimate an Apollo object hits Earth once every 250,000 years on average. A typical Apollo with a diameter of 1 km would strike with the power of a 100,000-megaton bomb and dig a crater more than 10 km in diameter. Astronomers estimate that there are about 1000 Apollos 1 km in size or larger, and the good news is that none of them are known to be on track to hit Earth in the foreseeable future. The bad news is that not all of these dangerous Apollo objects have been found yet.

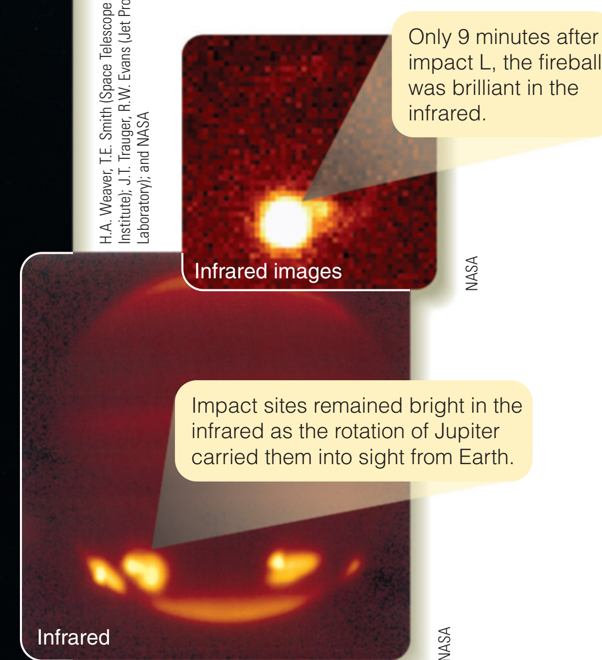
A significant danger lies in the more common smaller, yet still substantial, meteoroids, which are difficult to detect with current telescopes. For example, a rock designated 2011 MD, estimated to have been 15 m (50 ft) in diameter, about half the size of the Tunguska Impactor, passed only 12,000 km (7500 mi) from Earth's surface in June 2011. That is only twice the distance of humanity's geosynchronous communications and weather satellites, a very near miss indeed, yet DD45 was first spotted only three days before its closest approach. The future of our civilization on Earth may depend on our doing an increasingly careful job of tracking both large and small objects that cross our path with surprising frequency.



H.A. Weaver, T.E. Smith (Space Telescope Science Institute); J.T. Trauger, R.W. Evans (Jet Propulsion Laboratory); and NASA

NASA

© Cengage Learning 2014



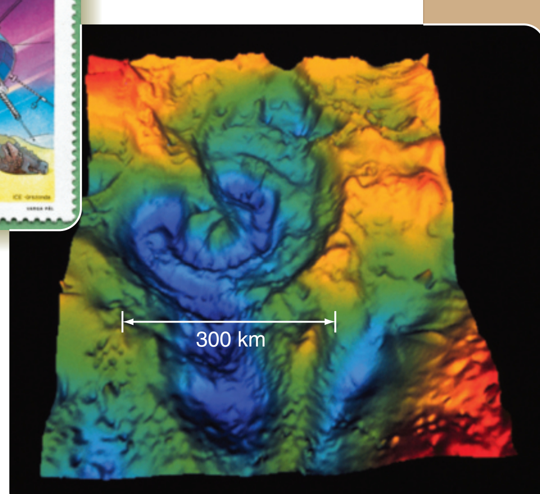
NASA

NASA

Figure 12-12

In 1992 Comet Shoemaker-Levy 9 passed within 1.3 planetary radii of Jupiter's center, well within its Roche limit, and tidal forces ripped the nucleus into more than 20 pieces. The fragmented pieces were as large as a few kilometers in diameter and spread into a long string of objects that looped away from Jupiter and then fell back to strike the planet, producing massive impacts over a period of six days in July 1994. Note that the impact flash in the upper right panel is larger than Earth.

Fair Use



Virgil L. Sharpton, University of Alaska, Fairbanks

© Cengage Learning 2014

Figure 12-13

The giant impact scar buried in Earth's crust near the village of Chicxulub in the northern Yucatán peninsula was formed about 65 million years ago by the impact of a large asteroid or comet. This gravity map shows the extent of the crater hidden below limestone deposited long after the impact.

What Are We? Sitting Ducks

Human civilization is spread out over Earth's surface and exposed to anything that falls out of the sky. Meteorites, asteroids, and comets bombard Earth, producing impacts that vary from dust settling gently on rooftops to disasters capable of destroying all life. In this case, the scientific evidence is conclusive and highly unwelcome.

Statistically we are quite safe. The chance that a major impact will occur during your lifetime is so small it is hard to estimate. But the consequences of such an impact are so severe that humanity should be preparing. One way to prepare is to find those objects that could hit us, map their orbits, and identify any that are dangerous.

What we do next isn't clear. Blowing up a dangerous asteroid in space might make a good movie, but converting one big projectile into a thousand small ones might not be very smart. Changing an asteroid's orbit could be difficult without a few decades' advance warning. Unlikely or not, large impacts demand consideration and preparation.

Throughout the universe there may be two kinds of inhabited worlds. On one type of world, intelligent creatures have developed ways to prevent asteroid and comet impacts from altering their climates and destroying their civilizations. But on other worlds, including Earth, intelligent races have not yet found ways to protect themselves. Some of those civilizations survive. Some don't.

Study and Review

Summary

- Review from Chapter 8: The term *meteoroid* refers to small solid particles orbiting in the solar system. The term *meteor* refers to a visible streak of light from a meteoroid heated and glowing as it enters Earth's atmosphere. The term *meteorite* refers to space material that has reached Earth's surface.
- **Iron meteorites (p. 239)** are mostly iron and nickel; when sliced open, polished, and etched, they show **Widmanstätten patterns (p. 240).** These reveal that the metal cooled from a molten state very slowly.
- **Stony meteorites (p. 239)** included **chondrites (p. 241),** which contain small, glassy particles called **chondrules (p. 241),** solidified droplets of once-molten material that formed in the solar nebula by an as yet unknown mechanism.
- **Stony-iron meteorites (p. 239)** are quite rare and, as the name implies, consist of a mixture of stony and metallic material.
- **Selection effects (p. 240)** cause iron meteorites to be the most common finds, even though stony meteorites are the most common falls.
- Stony meteorites that are rich in volatiles and carbon are called **carbonaceous chondrites (p. 240).** They are among the least modified meteorites. Some carbonaceous chondrites contain **CAIs (p. 241),** calcium-aluminum-rich inclusions, which are understood to be the very first solid particles to condense in the cooling solar nebula.
- An **achondrite (p. 241)** is a stony meteorite that contains no chondrules and no volatiles. Achondrites appear to have been melted after they formed and, in some cases, resemble solidified lavas.
- Evidence from the orbits of meteorites seen to fall, and from the composition of the meteorites, suggests that meteorites are fragments of asteroids. Other evidence, including orbital paths indicated by the **radiant (p. 242)** points of **meteor showers (p. 242),** indicates in contrast that the vast majority of meteors, including both meteors in meteor showers and isolated **sporadic meteors (p. 242),** appear to be low-density, fragile bits of debris from comets.
- Many meteorites appear to have formed as part of larger bodies that melted, differentiated, and cooled very slowly. Later these bodies were broken up, and fragments from the core became iron meteorites, fragments from the outer layers became stony meteorites, and fragments from intermediate layers became stony-irons.
- Asteroids are irregular in shape and heavily cratered from collisions. Their surfaces are covered by gray, pulverized rock, and some asteroids have such low densities they must be fragmented rubble piles.
- Most asteroids lie in a belt between Mars and Jupiter. Two other groups of asteroids called the **Trojan asteroids (p. 247)** are caught in the Lagrange points along Jupiter's orbit 60° ahead and 60° behind the planet. Asteroids that pass near Earth are called **Near-Earth Objects (NEOs) (p. 246).** NEOs with orbits that actually cross Earth's orbit and could eventually hit Earth are called **Apollo objects (p. 246). Centaurs (p. 247)** are asteroids that orbit among the planets of the outer solar system.
- The asteroids formed as rocky planetesimals between Mars and Jupiter, but Jupiter prevented them from forming a planet. Collisions have fragmented all but the largest of the asteroids. Most of the material inferred to have been originally in the asteroid belt has been gravitationally perturbed and swept up by the planets or tossed out of the solar system.
- **C-type (p. 245)** asteroids are more common in the outer asteroid belt where the solar nebula was cooler. They are darker and may be carbonaceous. **S-type (p. 245)** asteroids are the most common and may be the source of the most common kind of meteorites, the chondrites. Many **M-type (p. 245)** asteroids appear to have nickel-iron compositions and may be the cores of differentiated asteroids shattered by collisions.
- A visible comet is produced by a lump of ices and rock usually about 1 to 10 km in diameter, referred to as the comet nucleus. In long, elliptical orbits, the icy nucleus stays frozen until it nears the sun. Then, some of the ices vaporize and release dust and gas that is blown away to form a prominent head and tail.
- A comet's **gas tail (p. 250)** is ionized gas carried away by the solar wind. The **dust tail (p. 250)** is composed of solid debris released from the nucleus and blown outward by the pressure of sunlight. A comet's tail always points away from the sun, no matter in what direction the comet is moving.
- The **coma (p. 250),** or head, of a comet can be up to a million kilometers in diameter.

- Spacecraft flying past comets have revealed that they have very dark, rocky crusts and that jets of vapor and dust issue from active regions on the sunlit side.
- The low density of comet nuclei shows that they are irregular mixtures of ices and silicates, probably containing large voids. At least one comet nucleus has surface features showing the material has a surprising amount of strength.
- Comets are believed to have formed as icy planetesimals in the outer solar system, and some were ejected to form the **Oort cloud (p. 253).** Comets perturbed inward from the Oort cloud become long-period comets.
- Other icy bodies formed in the outer solar system and now make up the Kuiper belt beyond Neptune. Objects from the Kuiper belt that are perturbed into the inner solar system can become short-period comets.
- A major impact on Earth can trigger extinctions because it can substantially alter the entire world's climate.
- An impact at **Chicxulub (p. 255)** in Mexico's Yucatán region 65 million years ago appears to have triggered the extinction of 75 percent of the species then on Earth, including the dinosaurs.

Review Questions

1. What do Widmanstätten patterns indicate about the history of iron meteorites?
2. What do chondrules tell you about the history of chondrites?
3. Why are there no chondrules in achondritic meteorites?
4. Why do astronomers refer to carbonaceous chondrites as unmodified or "primitive" material?
5. How do observations of meteor showers reveal one of the sources of meteoroids?
6. How can most meteors be cometary if all meteorites are asteroidal?
7. Why do astronomers think the asteroids were never part of a full-sized planet?
8. What evidence indicates that the asteroids are mostly fragments of larger bodies?
9. What evidence indicates that some asteroids have differentiated?
10. What evidence indicates that some asteroids have had geologically active surfaces?
11. How is the composition of meteorites related to the formation and evolution of asteroids?
12. What is the difference between a gas tail and a dust tail? What does that tell you about the composition and origin of comets?
13. Why do short-period comets tend to have orbits near the plane of the solar system?
14. What are the hypotheses for how the bodies in the Kuiper belt and the Oort cloud formed?
15. **How Do We Know?** How would studying the chemical composition of only the brightest and most easily observed asteroids possibly yield misleading information about asteroids in general?

Discussion Questions

1. It has been suggested that humans may someday mine the asteroids for materials to build and supply space colonies. What kinds of materials could Earthlings get from asteroids? (*Hint:* What are S-, M-, and C-type asteroids made of?)
2. If cometary nuclei were heated during the formation of the solar system by internal radioactive decay rather than by solar radiation, how would comets differ from what is observed?
3. Do you think the government should spend money to find near-Earth asteroids? How serious is the risk?

Problems

1. Large meteorites are hardly slowed by Earth's atmosphere. Assuming the atmosphere is 100 km thick and that a large meteorite falls perpendicular to the surface, how long does it take to reach the ground? (*Hint:* Refer to the text for typical speeds of meteoroids.)
2. If a single asteroid 1 km in diameter were to be fragmented into meteoroids 1 m in diameter, how many would it yield? (*Hint:* The volume of a sphere $= \frac{4}{3}\pi r^3$.)
3. If a trillion (10^{12}) asteroids, each 1 km in diameter, were assembled into one body, how large would it be? (*Hint:* The volume of a sphere $= \frac{4}{3}\pi r^3$.) Compare that to the size of Earth.
4. The asteroid Vesta has a mass of 2.6×10^{20} kg and a radius of about 270 km. What is the circular orbit velocity at its surface? The fastest major league pitchers can throw a baseball at 105 mph (46.9 meters per second). Could they throw a baseball into orbit around Vesta? (*Hint:* Use the circular orbit velocity formula, Chapter 4.)
5. What is the maximum angular diameter of the largest asteroid, Ceres, as seen from Earth? Could Earth-based telescopes detect surface features? Could the *Hubble Space Telescope*? (*Hints:* Use the small-angle formula, Chapter 3. The angular resolution of Earth-based telescopes is about 1 arc second and of *Hubble* about 0.1 arc second. Ceres's average distance from the sun is 2.8 AU.)
6. What is the orbital period of Ceres? (*Hints:* Use Kepler's third law, Chapter 4; refer to the previous problem for Ceres's average distance from the sun.)
7. If the velocity of the solar wind is about 400 km/s and the visible tail of a comet is 1×10^8 km long, how much time does it take for a solar wind atom to travel from the nucleus to the end of the visible tail?
8. If you saw Comet Halley when it was 0.7 AU from Earth and it had a visible tail 5° long, how long was the tail in kilometers? Suppose that the tail was not perpendicular to your line of sight. Is your first answer too large or too small? (*Hint:* Use the small-angle formula, Chapter 3.)
9. What is the orbital period of a comet nucleus in the Oort cloud? What is its orbital velocity? (*Hints:* Use Kepler's third law, Chapter 4. The circumference of a circular orbit $= 2\pi r$. Refer to the text for typical Oort cloud object distances from the sun.)
10. The mass of an average comet's nucleus is about 10^{14} kg. If the Oort cloud contains 2×10^{12} comet nuclei, what is the mass of the cloud in Earth masses? Compare that with Jupiter's mass. (*Hint:* See Appendix Table A-10.)

Learning to Look

1. What do you see in the image below that tells you the size of planetesimals when the solar system was forming?

Lab photograph, R. Kempton, New England Meteoritical

2. Discuss the surface of the asteroid Mathilde, pictured below. What do you see that tells you something about the history of the asteroids?

NASA/JPL

3. What do you see in this image of the nucleus of Comet Borrelly that tells you how comets produce their comae and tails?

NASA

Great Debates

1. *Drawing Straws?* Drawing straws is a process used to randomly choose an individual for either an unfavorable task that lacks volunteers or a favorable task that has too many volunteers. Suppose Earth is subject to a large asteroid or comet impact that might destroy nearly all, if not all, of the human race. If a space pod, having a maximum capacity of 200 individuals, could dock with the Space Station for the worst of the impact, thereby saving humanity, should Earthlings be allowed to draw straws for a spot on the pod? If not, who in the world should be assigned a place in the pod and why?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers. Use at least one "impacting object" vocabulary word that can induce such an evacuation.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

2. *Should Earth Have an Evacuation Plan?* Asteroids 1300 ft (400 m) in diameter and larger have recently passed close to Earth, closer than the distance from the moon to Earth. For such a large asteroid, we should be able to track the asteroid and know in advance whether and where the asteroid will strike Earth. Currently, Earth does not have an evacuation or a disaster plan. Should the United Nations develop an evacuation or disaster plan?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers. Use at least one "impacting object" vocabulary word that can induce such an evacuation.
   b. What's the evidence? Find two examples of impacts to celestial bodies in our solar system to support your claim.
   c. Cite your sources.

3. *Act of God.* Some insurance companies consider meteor impacts to your car or home as an "act of God," which is a legal term for natural disaster events outside human control; often insurance will not cover them. Your car or homeowner's insurance policy might not mention acts of God and may not mention natural disasters like impacts, tsunamis, and flash floods. Do you know if you are covered for such an event? Should insurance companies cover damages caused by an act of God in their insurance policies?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find at least one example of a meteorite effect to a human.
   c. Cite your sources.

4. *Meteorite for Sale.* You are walking and see a meteor streak through the sky and land in a cornfield near you. You walk to the impact site and find small pieces of very heavy rock. As no other rocks are in the cornfield, you wisely deduce that the rocks have to be meteorites. You need a new car as your car is on its last legs (which is why you are walking). Do you pick up the meteorites, sell them, and buy a new Maserati? Do you share the profits with the owner of the farm? Or do you walk away because the meteorite is not yours? What do you do?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find at least one example of a meteorite being sold for profit.
   c. Cite your sources.

5. *Meteorite Alien Life.* You hear on the news that a NASA scientist has found evidence of life on a meteorite. She claims that the fossilized remains of alien microorganisms are not so different than those found on Earth. The find, if true, could mean that life began as the solar system began, and life on Earth did not begin on Earth. Your neighbor wants you to sign a letter to your congressperson to demand the government round up all meteorites like this one, put them on a rocket, and eject them back into space. After all, as your neighbor tells you, we don't want alien bacteria invading Earth and killing off our family, friends, and neighbors. What do you do? He is your neighbor, and you don't want to upset him. How do you handle the situation?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

6. *Family Vacation.* You parents poll you and your siblings and ask where you want to go on vacation this year. You are in astronomy class and learn about the Barringer Meteor Crater in Arizona and suggest an educational trip. Your siblings are younger and are squabbling over the Great Wall of China and the Great Barrier Reef. What is your argument to convince your parents that your suggestion is the best one?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

ENHANCED WebAssign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Origin of Asteroids and Meteorites
- Journey of a Comet
- Impacts on Earth

CengageNOW Virtual Astronomy Labs 2.0

Virtual ASTRONOMY LABORATORIES

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 9: Asteroids and Kuiper Belt Objects

Some astronomers have called asteroids "vermin of the sky." After taking a long photographic exposure of a distant star, nebula, or galaxy, a streak often would be found across the image. Such a streak, created by foreground solar system objects moving on the sky at a different pace than the sidereal rate driven by Earth's rotation, could be fatal to the intended investigation.

Some asteroids, including the first and largest one known, Ceres, were discovered accidentally, by astronomers intending to observe something else and finding a "star" not on their charts that moved, night by night or even hour by hour, against the stellar background. Many asteroids have been found by deliberate searches, and at least one was found by pointing a telescope using the lucky numbers on the back of the fortune in a Chinese cookie as celestial coordinates.

The name *asteroid* is in fact a misnomer. It means "starlike," which may describe how they appear superficially, but as you have learned in this chapter *planetoid* would be a more accurate term. The asteroids are planetlike material, understood to be remnant planetesimals or fragments of planetesimals left over from the formation of the planets, and they carry important clues to the history of the solar system.

Many people who have not taken a college astronomy course are nevertheless aware of the asteroid belt between Mars and Jupiter. But they usually don't know that the solar system contains a second such belt, beyond the orbit of Neptune, discovered by deliberate searches in either 1930 or 1992, depending on how the story is told. This second planetoid belt is named the Kuiper belt after Gerard Kuiper, a planetary astronomer who predicted its existence based on his understanding of how mass was distributed in the original disk of material from which the planets formed. The ambiguity in the year of discovery is due to the retroactive realization by astronomers that Pluto, found in 1930, is really just one of the larger (not the largest) objects in the Kuiper belt. The second object in that belt, still without a proper name but only the asteroid catalog designation 15760 1992 QB1, was found in 1992.

As of this writing, in summer 2012, there are more than a thousand Kuiper Belt objects (KBOs) known. Astronomers infer, based on the statistics of discoveries so far, that there are 100,000 KBOs with diameters of 100 km or more. That should be impressive; it means the Kuiper belt is probably 100 times more massive as the main asteroid belt.

Section 1 of Virtual Astronomy Lab 9, "Asteroids and Kuiper Belt Objects," lets you follow in the footsteps of planetary astronomers discovering these small but important members of our solar system. Section 2 shows you how to use Kepler's third law to make a first determination of the orbit radius of an object you have discovered. Section 3 explores the statistics of asteroid orbits that reveal the gravitational influence of their massive neighbor, Jupiter. Section 4 is a lesson in using the reflectivity, termed *albedo*, of asteroids to determine their surface temperatures and sizes. Section 5 links the history of the asteroids with the KBOs by considering the unexpected fact that both belts include binary objects. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

13

The Family of Stars

Guidepost

To discover the properties of stars, astronomers use telescopes and instruments such as photometers, cameras, and spectrographs in clever ways to learn the secrets hidden in starlight. The result is a family portrait of the stars. Knowing the distances to stars is the key to knowing most of their other properties, but measuring those distances is very difficult.

Here you will find answers to five important questions about stars:

- **How far away are the stars?**
- **How much energy do stars make?**
- **How do spectra of stars allow you to determine their temperatures?**
- **How big are stars?**
- **How much mass do stars contain?**

With this chapter, you leave our sun behind and begin your study of the billions of stars that dot the sky. In a sense, stars are the basic building blocks of the universe. If you hope to understand what the universe is, what our sun is, what our Earth is, and what we are, you need to understand stars.

Once you know how to find the basic properties of stars, you will be ready to trace the history of the stars from birth to death, a story that begins in the next chapter.

An image of the Jewel Box open star cluster, NGC 4755, located in the constellation Crux taken using the European Southern Observatory's (ESO's) Very Large Telescope (VLT) on Cerro Paranal in Chile. Several blue supergiant and one red supergiant stars can be seen. The cluster is about 6400 light years away. Characteristics of the stars in it provide evidence that the cluster is very young, only 16 million years old.

ESO/Y. Beletsky

[Love] is the star to every wandering bark,
Whose worth's unknown, although his
height be taken.

WILLIAM SHAKESPEARE
SONNET 116

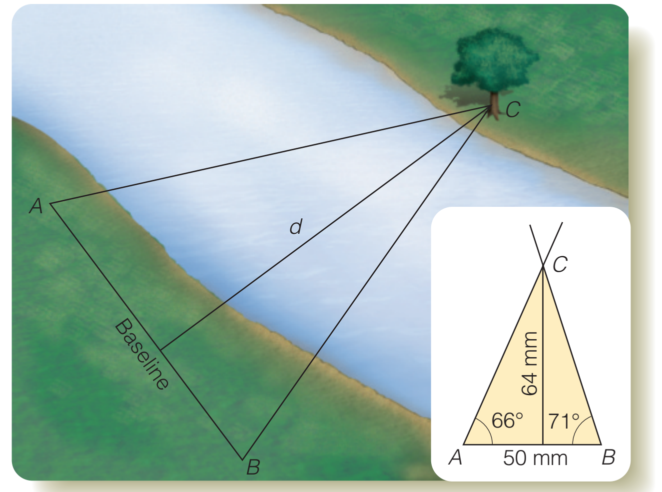


Figure 13-1

You can find the distance *d* across a river by measuring the length of the baseline and the angles *A* and *B,* and then constructing a scale drawing of the triangle.

SHAKESPEARE COMPARED LOVE to a star that can be seen easily and even used for guidance but whose real nature is utterly unknown. He lived at about the same time as Galileo and had no idea what stars actually are. To understand the history of the universe, the origin of Earth, and your place in the cosmos, you need to discover what people in Shakespeare's time did not know—the real nature of the stars. Unfortunately, it is quite difficult to find out what a star is like. When you look at a star even through a large telescope, you see only a point of light. Real understanding of stars requires careful analysis of starlight. This chapter concentrates on five goals: knowing how far away stars are, how much energy they emit, what their surface temperatures are, how big they are, and how much mass they contain. By the time you finish this chapter, you will know the family of stars well.

13-1 Star Distances

ALTHOUGH YOU WANT TO LEARN such things as the size and mass of stars, you immediately meet a detour. To find out almost anything about a star, you must know how far away it is.

Although knowing distances is crucial, it is also the most difficult measurement in astronomy. Astronomers have found a number of ways to estimate the distance to stars, but each of those ways ultimately depends on a direct geometrical method that is much like the method surveyors use to measure the distance across a river they cannot cross. You can begin by reviewing that method and then apply it to stars.

The Surveyor's Triangulation Method

To measure the distance to a landmark on the other side of a river, a team of surveyors begins by driving two stakes into the ground a known distance apart. The distance between the stakes is the *baseline* of the measurement. Using their surveying instruments, they sight the tree from the two ends of the baseline and measure the two angles on their side of the river, establishing a large triangle with corners marked by the two stakes and the tree (■ Figure 13-1).

Now that they know two angles of this large triangle and the length of the side between the angles, the surveyors can find the distance across the river by simple trigonometry. For example, if the baseline is 50 m long and the angles are 66° and 71°, the distance from the baseline to the tree works out to be 64 m. Another way to proceed would be to draw a small scale model of the big triangle. However you solve the problem, the point is that simple triangulation can reveal the distance across a river.

The more distant an object is, the longer the baseline you need to measure its distance accurately. You could use a baseline 50 m long to find the distance across a river; but to measure the distance to a mountain on the horizon, you might need a baseline 1000 m long. Great distances require long baselines.

The Astronomer's Triangulation Method

To find the distance to a star, you must use a very long baseline, the diameter of Earth's orbit. If you take a photograph of a nearby star and then wait six months, Earth will have moved halfway around its orbit. You can then take another photograph of the star from the other side of Earth's orbit. In this example your baseline equals the diameter of Earth's orbit, or 2 AU, and lines of sight to the star from the two observing positions outline a long thin triangle (■ Figure 13-2).

You then have two photographs of the same part of the sky taken from slightly different locations in space. When you examine the photographs, you will discover that the nearby star is not in exactly the same place in the two photographs. This shift in the position of the star is called *parallax,* the apparent change in the position of an object due to a change in the location of the observer. In Chapter 4 (page 48), you saw an everyday example. Your thumb, held at arm's length, appears to shift position against a distant background when you look first with one eye and then with the other. In that case, the baseline is the distance between your eyes, and the parallax is the angle through which

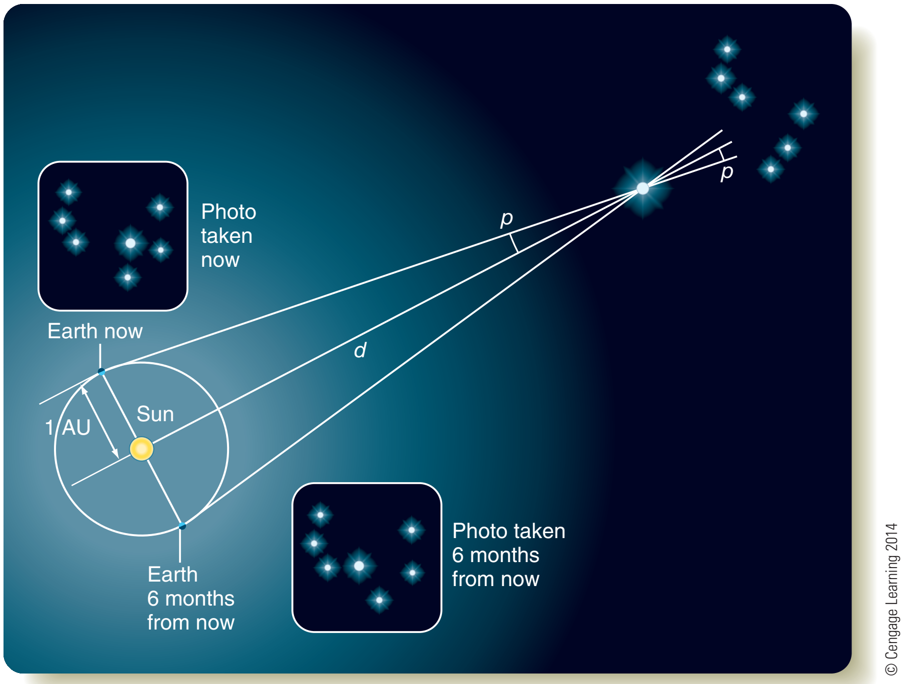


Figure 13-2

You can measure the parallax of a nearby star by photographing it from two points around Earth's orbit. For example, you might photograph it now and again in six months.

your thumb appears to move when you change eyes. The farther away you hold your thumb, the smaller its parallax.

Because the stars are so distant, their parallaxes are very small angles, usually expressed in arc seconds. Astronomers conventionally call **stellar parallax *(p)*** the shift of the star observed across a 1 AU (not 2 AU) baseline, as shown in Figure 13-2.

Astronomers measure the parallax, and surveyors measure the angles at the ends of the baseline, but both measurements reveal the same thing—the shape and size of the triangle and thus the distance to the object in question.

The distances to stars are so large that astronomers have defined a special unit of distance, the **parsec (pc),** for use in distance calculations (**Reasoning with Numbers 13-1**). A parsec is defined as the distance to an imaginary star with a parallax of 1 arc second; the word parsec was created by combining parallax and arc second. One parsec equals 2.06×10^5 AU, or 3.26 ly. The parsec unit is used routinely by astronomers because it simplifies calculation of distance from observed parallax shifts. However, there are instances when units of light-years are also convenient. The chapters that follow this one use either parsecs or light-years as convenience and custom dictate.

Reasoning with Numbers | 13-1

Parallax and Distance

To find the distance to a star from its measured parallax, astronomers use the same calculation you have already seen in the small-angle formula (Reasoning with Numbers 3-1). Imagine that you observe our solar system from the star. Figure 13-2 shows that the angular separation you would measure between the sun and Earth equals the star's parallax *p.* Recall that the small-angle formula relates an object's angular diameter, its linear diameter, and its distance. In this case, the angular diameter is the parallax angle *p* and the linear diameter, the base of the triangle, is 1 AU. Then the small-angle formula, rearranged slightly, tells you that the distance *d* to the star in AU is equal to 2.06×10^5 divided by the parallax in arc seconds:

$$d(AU) = \frac{2.06 \times 10^5}{p}$$

The constant 2.06×10^5 is a conversion factor, the number of arc seconds in a radian.

Because the parallaxes of even the nearest stars are less than 1 arc second, the distances in AU are inconveniently large numbers. To keep the numbers manageable, astronomers have defined the parsec as their primary unit of distance in a way that simplifies the arithmetic. One parsec equals 2.06×10^5 AU, so the equation becomes:

$$d(pc) = \frac{1}{p}$$

Thus, a parsec is the distance to an imaginary star whose parallax is 1 arc second.

Example: The star Altair has a parallax of 0.194 arc second. How far away is it? **Solution:** The distance in parsecs equals 1 divided by 0.194, or 5.16 pc:

$$d(pc) = \frac{1}{0.194} = 5.16 \text{ pc}$$

One parsec equals about 3.26 light-years, so Altair is 16.8 ly away.

Measuring the parallax p is very difficult because it is such a small angle. The star nearest the sun is one of your Favorite Stars, Alpha Centauri. It has a parallax of only 0.747 arc second, and the more distant stars have even smaller parallaxes. To see how small these angles are, hold a piece of paper edgewise at arm's length. The thickness of the paper covers an angle of about 30 arc seconds.

The blurring caused by Earth's atmosphere smears star images and makes them about 1 arc second in diameter even at a good observatory site, and that makes it difficult to measure parallax from Earth. Even when astronomers average together many observations, they cannot measure parallax from an observatory on Earth with an uncertainty smaller than about 0.002 arc second. Therefore, if you measure a parallax of 0.02 arc second from Earth, the uncertainty is about 10 percent. Ten percent is about the largest uncertainty in a parallax measurement that astronomers are willing to tolerate, so ground-based astronomers have not been able to measure the distance to stars more distant than about 50 pc. Since the first stellar parallax was measured in 1838, astronomers using ground-based telescopes have been able to measure accurate parallaxes for only about 10,000 stars.

In 1989, the European Space Agency launched the satellite *Hipparcos* to measure stellar parallaxes from orbit above the blurring effects of Earth's atmosphere. The satellite observed for four years, and the data were used to produce two parallax catalogs in 1997. One catalog contains 120,000 stars with parallaxes 20 times more precise than ground-based measurements. The other catalog contains over a million stars with parallaxes as accurate as ground-based parallaxes. The *Hipparcos* data have given astronomers new insights into the nature of stars.

The European Space Agency expects to launch the *Gaia* mission in a few years. It will be able to measure the parallaxes of over a billion stars as faint as apparent visual magnitude +20. This will allow the first real 3-dimensional map of our Galaxy.

13-2 Apparent Brightness, Intrinsic Brightness, and Luminosity

If you see a light on a dark highway, it is hard to tell how powerful it really is. It could be the brilliant headlight on a distant truck or the dim headlight on a nearby bicycle (■ Figure 13-3). How bright an object appears depends not only on how much light it emits but also on its distance.

A sixth-magnitude star just visible to your eye looks faint, but its apparent magnitude doesn't tell you how luminous it really is. Now that you know how to find the distance to stars, you can use those distances to figure out the **intrinsic brightness** of the stars. *Intrinsic* means "belonging to the thing," so the intrinsic brightness of a star refers to the total amount of light the star emits.

Brightness and Distance

When you look at a bright light, your eyes respond to the visual wavelength photons falling on your eye's retina. The apparent brightness you perceive is related to the flux of energy entering your eye. **Flux** is the energy in joules (J) per second falling on 1 square meter. Recall that a joule is about as much energy as is released when an apple falls from a table onto the floor. One joule per second is one watt, a common unit of energy consumption used, for example, to rate light bulbs.

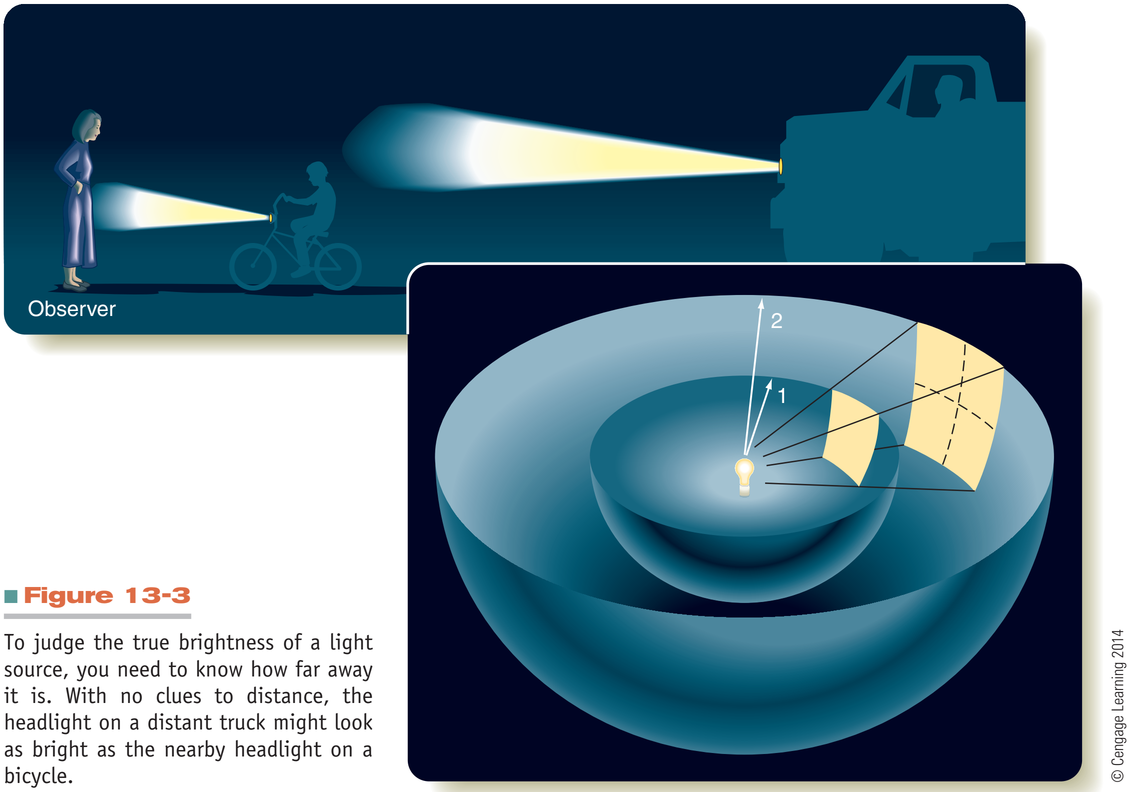
The apparent brightness of a light source is determined by the inverse square law. You first encountered the inverse square relation in Chapter 4, where it was applied to the strength of gravity. If you double the distance to a light source, its brightness falls by a factor of 2^2, or 4 times. If you triple the distance, the brightness falls by a factor of 3^2, which equals 9 times. In other words, the flux of light reaching you is inversely proportional to the square of your distance from the source.

You can see that if you know both the apparent magnitude of a star (expressing the flux received on Earth from it) and its distance, you can use the inverse square law to correct for distance and learn the intrinsic brightness of the star. Astronomers do that using a special kind of magnitude scale described in the next section.

Absolute Visual Magnitude

If all stars were the same distance from Earth, you could compare one with another and decide which was emitting more light and which less. Of course, the stars are scattered at different distances, and you can't shove them around to line them up for comparison. If, however, you know the distance to a star, you can use the inverse square relation to calculate the brightness the star would have at some standard distance. Astronomers have adopted 10 pc as the standard distance and refer to the **absolute visual magnitude (M_V)** of a star as the apparent visual magnitude it would have if it were 10 pc away. The absolute visual magnitude is therefore an expression of the intrinsic brightness of the star.

To find a star's absolute visual magnitude, you begin by measuring its apparent visual magnitude (in other words, its brightness), a relatively easy task. Then, you need the distance to the star. If the star is nearby, you can measure its parallax and from that calculate the distance. Once you know the distance, you can use a simple formula to correct the apparent visual magnitude for the distance and find the absolute visual magnitude (**Reasoning with Numbers 13-2**).



■ **Figure 13-3**

To judge the true brightness of a light source, you need to know how far away it is. With no clues to distance, the headlight on a distant truck might look as bright as the nearby headlight on a bicycle.

Reasoning with Numbers | 13-2

Absolute Magnitude and Distance

Apparent visual magnitude tells you how bright a star looks (see Reasoning with Numbers 2-1), but absolute visual magnitude tells you how luminous the star really is. The absolute visual magnitude M_V of a star is the apparent visual magnitude the star would have if it were 10 pc away. If you know a star's apparent visual magnitude and its distance, you can calculate its absolute visual magnitude. The **magnitude-distance formula** that allows this calculation relates apparent visual magnitude m_V, distance in parsecs *d*, and absolute visual magnitude M_V:

$$m_V - M_V = -5 + 5 \log(d)$$

The expression "log" means logarithm to the base 10.

Sometimes it is convenient to rearrange the equation and write it in this form:

$$d = 10^{(m_V - M_V + 5)/5}$$

It is the same equation, so you can use whichever form is most convenient in a given problem. If you know the distance, the first form of the equation is convenient, but if you are trying to find the distance, the second form of the equation is best.

Example: Favorite Star Polaris is 133 pc from Earth and has an apparent magnitude of 2.0. What is its absolute visual magnitude? **Solution:** A pocket calculator tells you that log(133) equals 2.12, so you substitute into the first equation to get

$$2.0 - M_V = -5 + 5(2.12)$$

Solving for M_V tells you that the absolute visual magnitude of Polaris is −3.6. If it were only 10 pc from Earth, it would dominate the night sky.

The subscript "V" in the symbol for absolute visual magnitude reminds you that it is a visual magnitude including only the wavelengths of light you can see. Other magnitude systems are based on other parts of the electromagnetic spectrum, such as the infrared and ultraviolet.

How does the sun stack up against other stars? The sun is tremendously bright in the sky, but it is also very nearby. Its absolute visual magnitude is just 4.83. If the sun were only 10 pc from Earth (about 33 ly, not a great distance in astronomy), it would look no brighter than the faintest star in the handle of the Little Dipper.

The intrinsically brightest stars known have absolute visual magnitudes of about −8, which means that such a star 10 pc from Earth would be nearly as bright as the moon. Such stars have intrinsic brightness 13 magnitudes brighter than the sun, which means they are emitting over 100,000 times more light than the sun. In contrast, the intrinsically faintest stars have absolute visual magnitudes of +15 or fainter. They are ten magnitudes fainter than the sun, meaning they are emitting 10,000 times less light at visible wavelengths than the sun. You now have some real insight into the wide range of star characteristics.

Luminosity

The **luminosity** ***(L)*** of a star is the total energy the star radiates in one second. Hot stars emit a great deal of ultraviolet light that you can't see, and cool stars emit mostly infrared light. Absolute visual magnitude includes only visible radiation, so astronomers must make a correction, sometimes quite large, to account for the invisible energy. Then they can calculate the total luminosity of the star from its absolute magnitude.

Astronomers often express luminosities in solar units, meaning that they write $2.5\ L_{\odot}$ to represent a star that has 2.5 times the luminosity of the sun. To find the luminosity of a star in joules per second, you can just multiply by the luminosity of the sun in those units, 3.8×10^{26} J/s. For example, Favorite Star Aldebaran has a luminosity of about $150\ L_{\odot}$, which corresponds to about 5.7×10^{28} J/s.

The most luminous stars emit at least a billion times more energy per second than the least luminous. Clearly, the family of stars contains some interesting characters.

SCIENTIFIC ARGUMENT

How can two stars look the same in the sky but have dramatically different intrinsic brightness?

You can answer this question by building a scientific argument that relates three factors: the apparent brightness of a star, its intrinsic brightness, and its distance. The further away a star is, the fainter it looks—that is just the inverse square law. Favorite Stars Vega and Rigel have the same apparent visual magnitude, which means your eyes receive the same amount of light from them. But parallax observations from the *Hipparcos* satellite reveal that Rigel is 31 times further away than Vega, so Rigel is much more luminous than Vega.

Distance is often the key to determining the luminosities of stars, but temperature can also be important. Build a scientific argument to answer the following: **Why must astronomers make a correction in converting the absolute visual magnitudes of very hot or very cool stars into luminosities?**

13-3 Star Spectra

Observations of spectral lines give you information about what types of atoms are in the atmospheres of the sun, planets, and stars, as you learned in Chapter 6. In the late 1800s and early 1900s, when astronomers were making the first careful studies of stellar spectra, the big differences observed between spectra of different stars were thought to show that stars have a wide range of compositions.

You learned in Chapter 7 that in the 1920s Cecilia Payne made use of information from atomic physics and the new field of quantum mechanics to reinterpret the spectra of stars. Payne's calculations showed that over 90 percent of the atoms in the sun and other stars must be hydrogen and most of the rest are helium (look back to Table 7-1, page 123). Payne discovered that: (1) the chemical composition of the sun is like the composition of other stars, (2) spectra actually provide information mostly about the temperatures of stars, and (3) stars with similar spectra must have similar temperatures.

The Balmer Thermometer

Look back to "Atomic Spectra" on pages 110–111 in Chapter 6 and review how spectral absorption lines are produced in stellar atmospheres. The light that forms a spectrum comes from the gases of the surface and the atmosphere, so the spectrum is direct information only about those outer layers.

One of the main methods Payne first developed for using spectra to determine star temperatures is now called the Balmer thermometer. Recall that astronomers use the Kelvin temperature scale when referring to stellar temperatures. These temperatures range from about 40,000 K to about 2500 K; compare these extremes with the surface temperature of the sun, which is about 5800 K. From the information about blackbody radiation and Wien's law in Chapter 6, you already know how to estimate a star's temperature by using its color, but the spectral lines of hydrogen at wavelengths visible to the human eye, called the

Balmer lines, combined with a few other spectral lines, give you much greater precision in estimating stellar temperatures.

The Balmer thermometer works because the strength of the Balmer lines depends on the temperature of the star's surface layers. Both hot and cool stars have weak Balmer lines, but medium-temperature stars have strong Balmer lines. That is because the Balmer absorption lines are produced only by atoms with electrons in the second energy level.

If a star is cool, there are few violent collisions between atoms to excite the electrons, so the electrons of most atoms are in the ground state, not the second level. Electrons in the ground state can't absorb photons in the Balmer series. As a result, you should expect to find weak Balmer absorption lines in the spectra of cool stars.

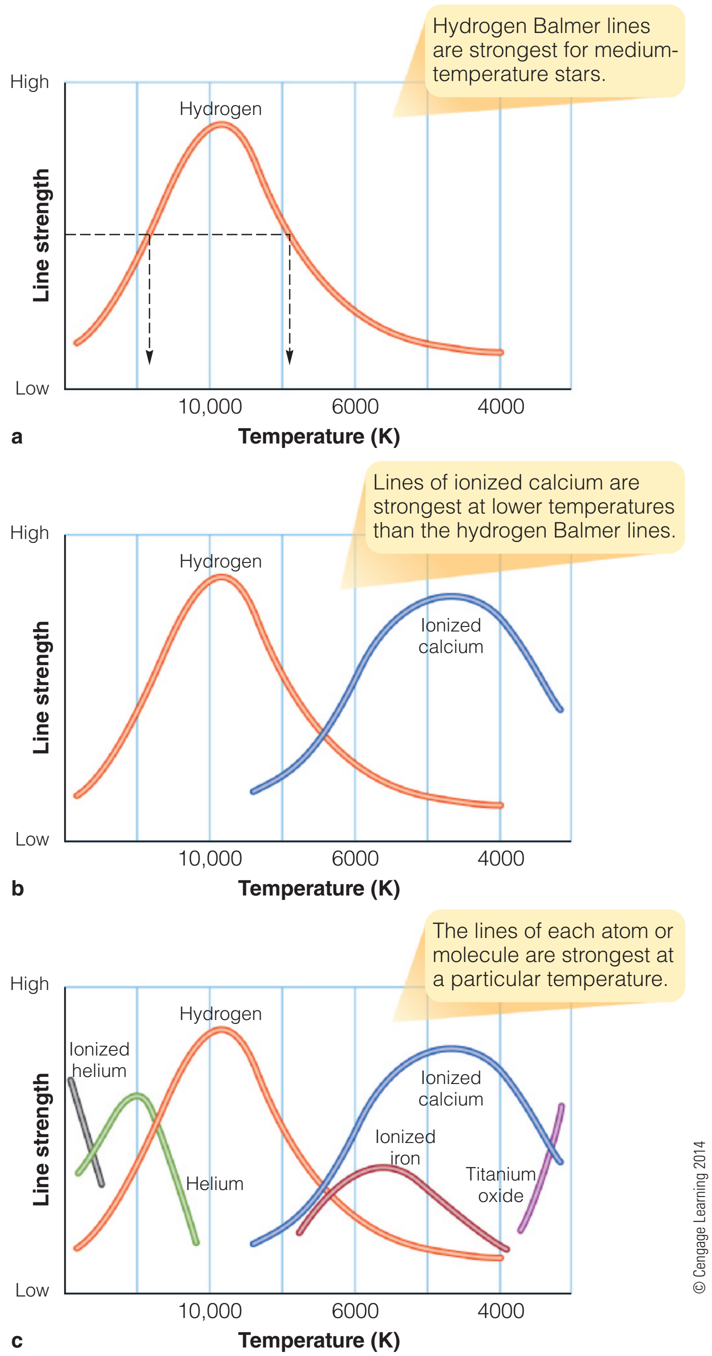
In the surface layers of hot stars, on the other hand, there are many violent collisions between atoms. These collisions can excite electrons to high energy levels or even ionize some atoms by knocking electrons completely out of them. In that situation also, there are few hydrogen atoms with their electrons in the second orbit to form Balmer absorption lines. Hot stars, like cool stars, have weak Balmer absorption lines.

In stars of an intermediate temperature, roughly 10,000 K, the collisions are just right to excite large numbers of electrons into the second energy level. Hydrogen gas at that temperature absorbs photons very well at the wavelengths of the Balmer lines and produces strong (dark) spectral lines.

Theoretical calculations can predict how strong the Balmer lines should be for stars of various temperatures. Such calculations are the key to finding temperatures from stellar spectra. The curve in ■ Figure 13-4a shows the strength of the Balmer lines for various stellar temperatures. You can see from the graph that a star with Balmer lines of a certain strength might have either of two temperatures, one high and one low. How do you know which is right? You have to examine spectral lines of substances other than hydrogen to be sure you have the correct temperature.

Temperature has a similar effect on the spectral lines of other elements, but the temperature at which the lines reach their maximum strength differs for each element (Figure 13-4b). If you add a number of chemical elements to your graph, you will have a powerful aid for finding the temperatures of stars (Figure 13-4c). For example, if you find medium-strength Balmer lines and strong helium lines in a star's spectrum, you can conclude that it has a temperature of about 20,000 K. But if the star has weak hydrogen lines and strong lines of ionized iron, you would assign it a temperature of about 5800 K, similar to that of the sun.

The spectra of stars cooler than about 3500 K contain dark bands produced by molecules such as titanium oxide (TiO). Because of their structure, molecules can absorb photons at



■ **Figure 13-4**

The strength of spectral lines can tell you the temperature of a star. (a) Balmer hydrogen lines alone are not enough because they give two answers. In other words, Balmer lines of a certain strength could be produced by a hotter star or a cooler star. (b) Adding another atom to the diagram helps, and (c) adding many atoms and molecules to the diagram creates a precise aid to determine the temperatures of stars.

many wavelengths, producing numerous, closely spaced spectral lines that blend together to form bands. These molecular bands appear in the spectra of only the coolest stars because, as mentioned before, molecules in cool stars are not subject to the violent collisions that would break them apart in hotter stars.

Temperature Spectral Classification

During the 1890s astronomers at Harvard Observatory invented the first widely used system for classifying stellar spectra. One of those scientists, Annie J. Cannon, personally inspected the spectra of over 250,000 stars. Spectra were first classified into groups labeled A through Q, but some of those groups were later dropped, merged with others, or reordered. The final classification scheme includes seven major temperature **spectral classes,** or **types,** still used today: O, B, A, F, G, K, M.*

This sequence of spectral types, called the **spectral sequence,** is a temperature sequence. The O stars are the hottest, and temperature decreases along the sequence to the M stars, the coolest. For better definition, astronomers divide each spectral class into ten subclasses. For example, spectral class A consists of the subclasses A0, A1, A2, . . . A8, A9. Next come F0, F1, F2, and so on. This finer division gives a star's temperature to a precision of about 5 percent. The sun, for example, is not just a G star, but a G2 star. ■ Table 13-1 breaks down some of the information contained in Figure 13-4c and presents it according to spectral class. For example, if a star has weak Balmer lines and lines of ionized helium, it must be an O star.

Thirteen stellar spectra are arranged in ■ Figure 13-5 from the hottest at the top to the coolest at the bottom. You can easily see in those spectra how the strength of spectral lines depends on temperature; the Balmer thermometer you have just learned about is especially obvious. The hydrogen Balmer lines are strongest in A stars that have middle-range temperatures, weak in hotter stars (O and B), and weak in cooler stars (F through M)

Although these color images of spectra are attractive, astronomers today normally do not work with spectra in the form of images. Rather, as you learned in Chapter 6, spectra are usually displayed as graphs of intensity versus wavelength that show dark absorption lines as dips in the graph (■ Figure 13-6). Such graphs allow more detailed analysis than photographs. Notice, for example, that the overall curves are segments of blackbody curves with spectral lines superimposed. The wavelength of maximum intensity is in the infrared for the coolest stars and in the ultraviolet for the hottest stars. Look carefully at these graphs, and you can see that helium lines are visible only in the spectra of the hottest classes and titanium oxide bands only in the coolest. Two lines of ionized calcium increase in strength from A to K and then decrease from K through M. Because the strengths of these spectral lines depend on temperature, it requires only a few moments to study a star's spectrum and determine its temperature.

Now you can learn something new about your Favorite Stars. Sirius, brilliant in the winter sky, is an A1 star; and Vega, bright overhead in the summer sky, is an A0 star. They have nearly the same temperature and color, and both have strong Balmer lines in their spectra. The bright red star in Orion is Betelgeuse, a cool M2 star, but blue-white Rigel is a hot B8 star. Polaris, the North Star, is an F8 star a bit hotter than our sun, and Alpha Centauri, the closest star to the sun, is a G2 star just like the sun.

The study of spectral types is more than a century old, but astronomers continue to discover and define new types. The **L dwarfs,** found in 1998, are cooler and fainter than M stars. They are understood to be objects smaller than stars but larger than planets called **brown dwarfs,** which you will learn more about in a later chapter. The spectra of M stars contain bands produced by metal oxides such as titanium oxide (TiO), but L dwarf spectra contain bands produced by molecules such as iron hydride (FeH). The **T dwarfs** are an even cooler and fainter type of brown dwarf than L dwarfs. Their spectra show absorption by methane (CH_4) and water vapor (■ Figure 13-7). In 2011, astronomers using infrared space telescopes, large ground-based telescopes, and highly sensitive infrared detectors discovered a class of objects with temperatures below 500 K that are labeled **Y dwarfs.**

*Generations of astronomy students have remembered the spectral sequence using the mnemonic "Oh, Be A Fine Girl (Guy), Kiss Me." Recent suggestions from students include "Oh Boy, An F Grade Kills Me" and "Only Bad Astronomers Forget Generally Known Mnemonics."

Table 13-1 | Temperature Spectral Classes

Spectral Class	Approximate Temperature (K)	Hydrogen Balmer Lines	Other Spectral Features	Naked-Eye Example
O	40,000	Weak	Ionized helium	λ Orionis (O8)
B	15,000	Medium	Neutral helium	Achernar (B3)
A	8500	Strong	Ionized calcium weak	Sirius (A1)
F	6600	Medium	Ionized calcium weak	Canopus (F0)
G	5500	Weak	Ionized calcium medium	Sun (G2)
K	4100	Very weak	Ionized calcium strong	Arcturus (K2)
M	3000	Very weak	TiO strong	Betelgeuse (M2)

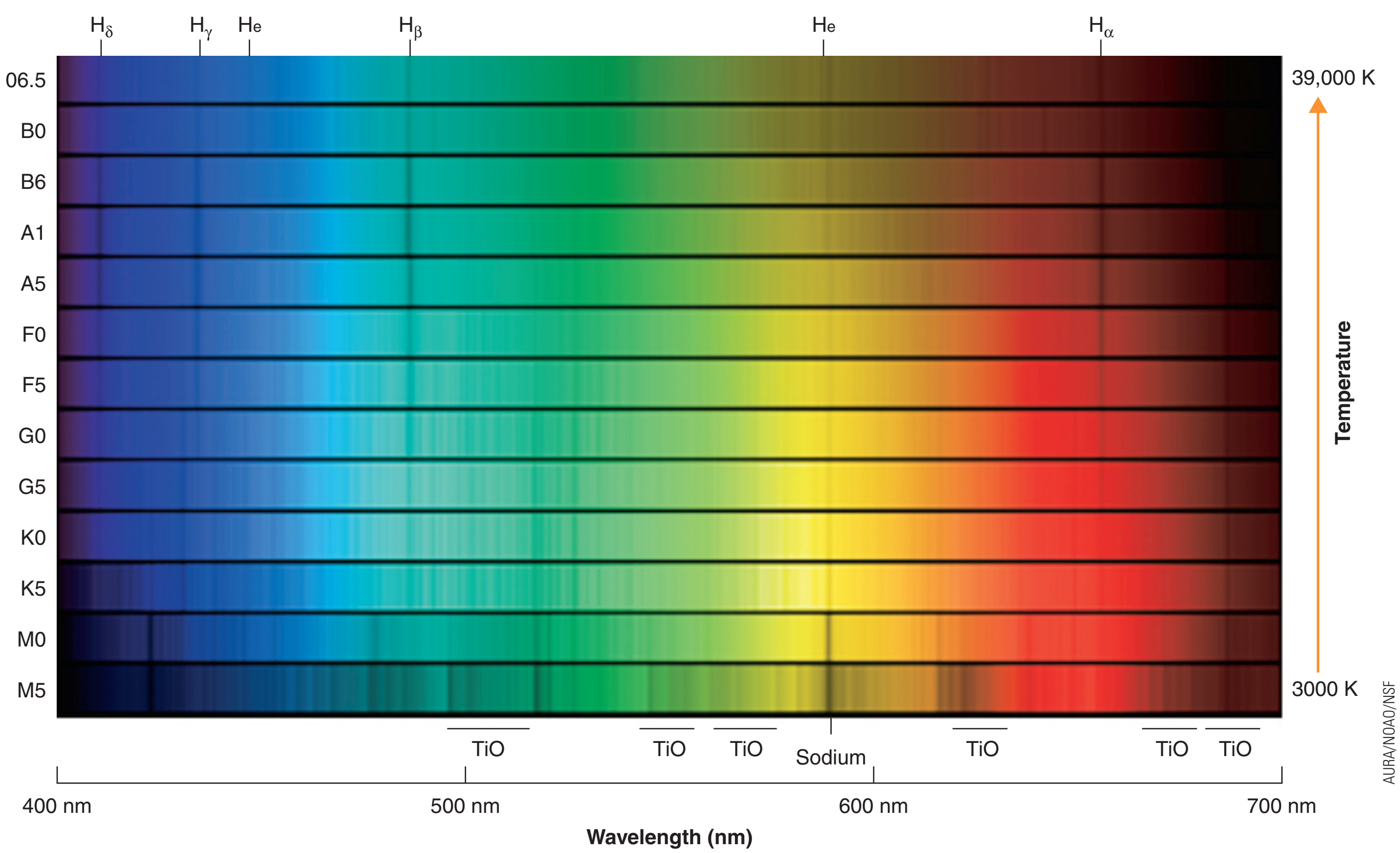


Figure 13-5

Color photographs of stellar spectra ranging from hot O stars at the top to cool M stars at the bottom. The Balmer lines of hydrogen are strongest at spectral type A0, but the two closely spaced lines of sodium in the yellow are strongest for very cool stars. Helium lines appear only in the spectra of the hottest stars. Notice that the helium line visible in the top spectrum has nearly but not exactly the same wavelength as the sodium lines visible in cooler stars. Bands produced by the molecule titanium oxide are strong in the spectra of the coolest stars.

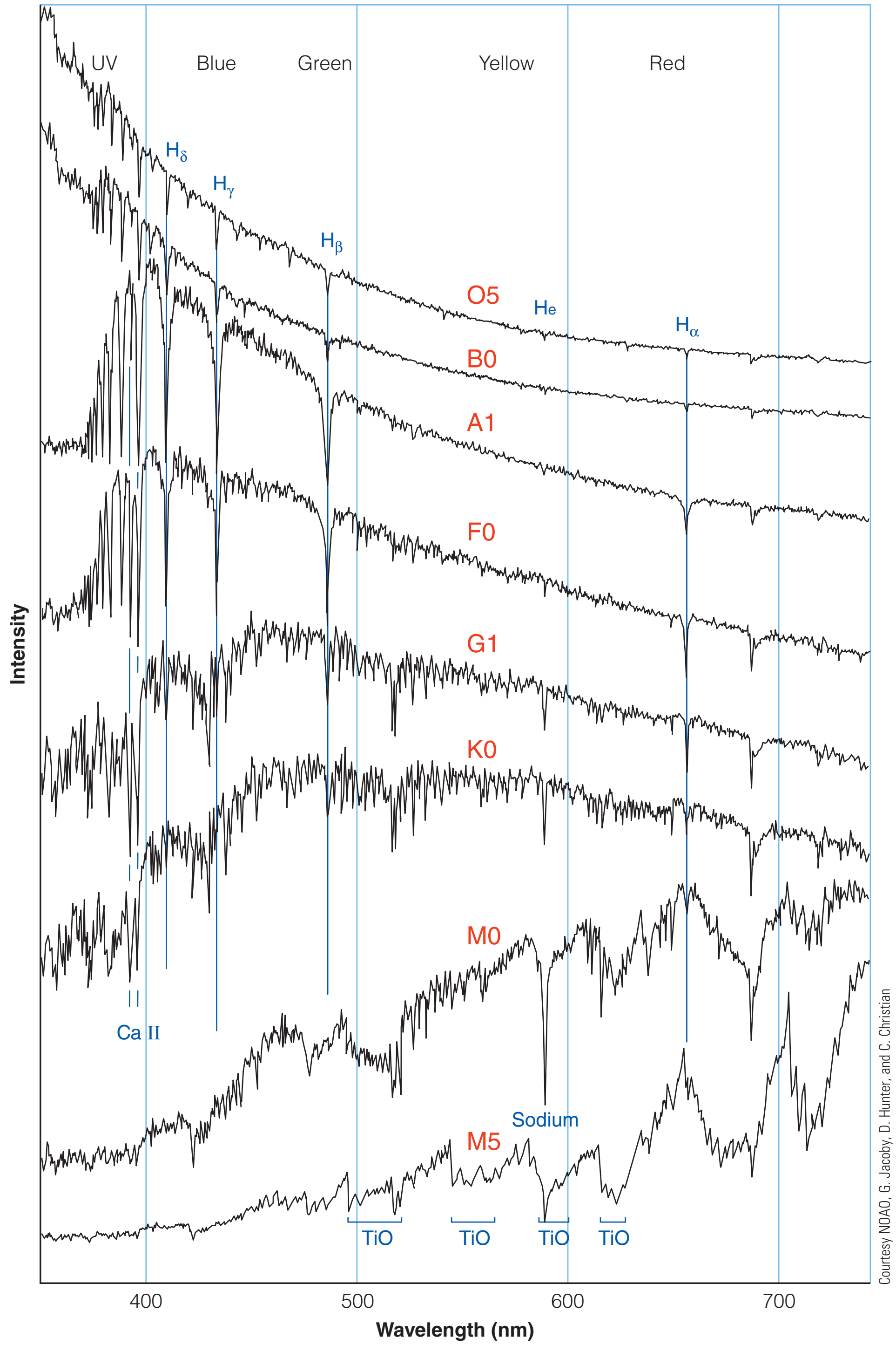


Figure 13-6

Modern digital spectra are usually represented as graphs of intensity versus wavelength, with dark absorption lines appearing as sharp dips in the curves. The hottest stars are at the top and the coolest at the bottom. Hydrogen Balmer lines are strongest at about A0, while lines of ionized calcium (Ca II) are strong in K stars. Titanium oxide (TiO) bands are strongest in the coolest stars. Compare these spectra with Figures 13-4c and 13-5.

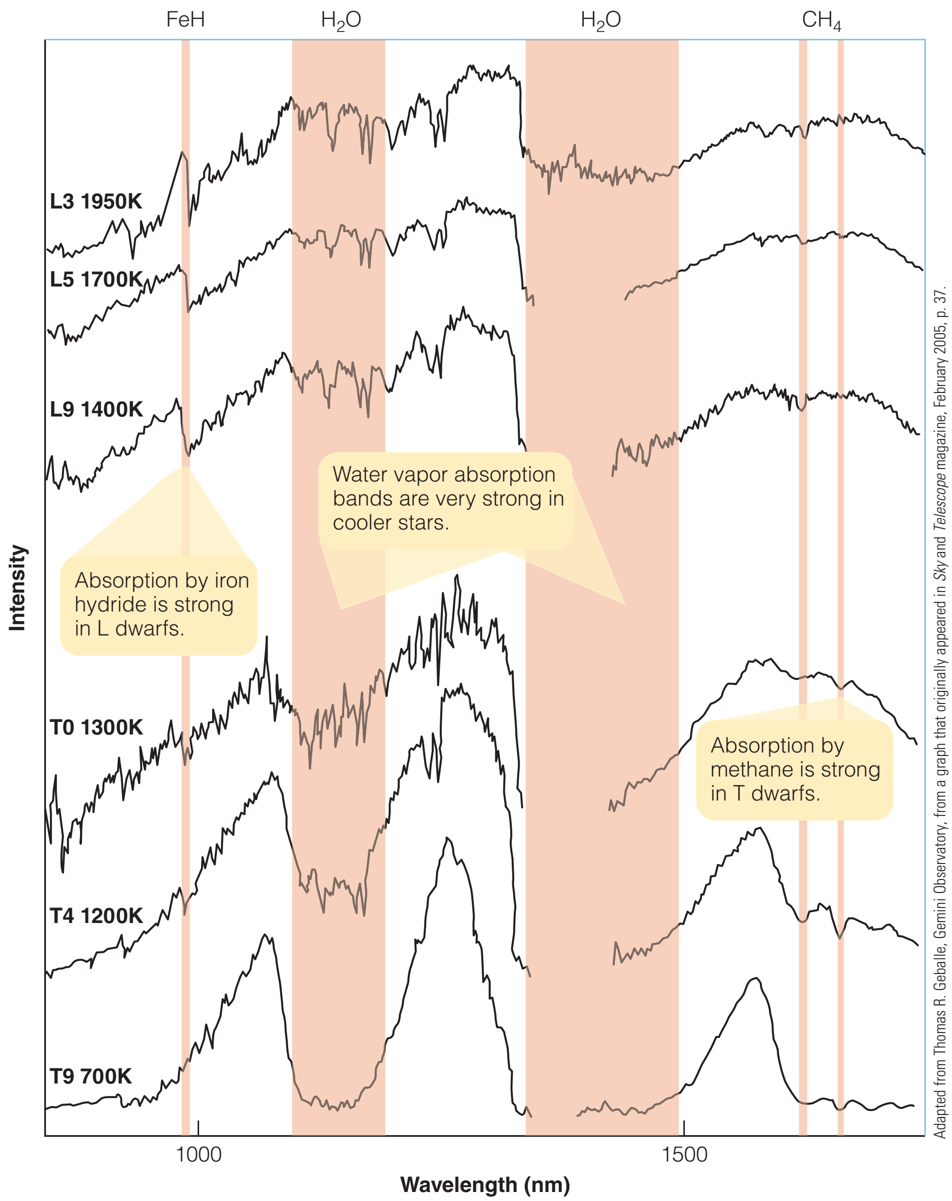


Figure 13-7

These six infrared spectra show the dramatic differences between L dwarfs and T dwarfs. Spectra of M stars show titanium oxide bands (TiO), but L and T dwarfs are so cool that TiO molecules do not form. Other molecules, such as iron hydride (FeH), water (H_2O), and methane (CH_4), can form in these very cool objects.

13-4 Star Sizes

Now that you know the luminosities of stars, you are ready to find their sizes, usually expressed as radii or diameters. Recall that astronomers can't see stars as disks through astronomical telescopes. The diameters of a few stars have been measured and surface features on a very small number of stars, including Favorite Star Betelgueuse, have been distinguished using the technique of interferometry (look back to Chapter 5), but essentially all stars look like points of light, no matter how big the telescope. Nevertheless, there is a straightforward way to find the sizes of stars. If you know a star's temperature and luminosity, you can determine its radius (**Reasoning with Numbers 13-3**). That relationship will introduce you to the most important diagram in astronomy, which sorts stars by temperature, luminosity and size, and in later chapters will help you learn about the life cycles of stars.

Luminosity, Radius, and Temperature

To use the luminosity and temperature of a star to find its size, you first need to understand the two factors that affect a star's luminosity: surface area and temperature. You can eat dinner by candlelight because a candle flame has a small surface area. Although the flame is very hot, it cannot radiate much heat; it has a low luminosity. However, if the candle flame were 12 ft tall, it would have a very large surface area from which to radiate, and, although it might be no hotter than a normal candle flame, its luminosity would drive you from the table (Figure 13-8).

In a similar way, a hot star may not be very luminous if it has a small surface area, but it could be highly luminous if it were larger and had a larger surface area from which to radiate. On the other hand, even a cool star could be luminous if it had a large surface area. Because of this dependence on both temperature and surface area, you need to separate their effects to find the sizes of stars.

The H–R Diagram

The **Hertzsprung–Russell (H–R) diagram,** named after its originators, Netherlands astronomer Ejnar Hertzsprung and U.S. astronomer Henry Norris Russell, is a graph that separates the effects of temperature and surface area on stellar luminosities and enables astronomers to sort and classify stars according to their sizes. Before you explore the details of the H–R diagram, try looking at a similar diagram you might use to describe and sort automobiles.

Reasoning with Numbers | 13-3

Luminosity, Radius, and Temperature

The luminosity L of a star depends on two things—its size and its temperature. If the star has a large surface area from which to radiate, it can radiate a great deal. Recall from the discussion of blackbody radiation in Reasoning with Numbers 6-1 that the amount of energy emitted per second from each square meter of the star's surface is σT^4. Thus, the star's luminosity can be written as its surface area in square meters times the amount it radiates from each square meter:

$$L = (\text{surface area}) \times \sigma T^4$$

Because a star is a sphere, you can use the formula: surface area $= 4\pi R^2$. Then the luminosity is:

$$L = 4\pi R^2 \sigma T^4$$

This may seem complicated, but if you express luminosity, radius, and temperature in proportion to the sun, you get a simpler form:*

$$\frac{L}{L_\odot} = \left(\frac{R}{R_\odot}\right)^2 \left(\frac{T}{T_\odot}\right)^4$$

Example A: Suppose a star has 10 times the sun's radius but only half the temperature. How luminous is it? **Solution:**

$$\frac{L}{L_\odot} = \left(\frac{10}{1}\right)^2 \left(\frac{1}{2}\right)^4 = \frac{100}{1} \times \frac{1}{16} = 6.25$$

This star therefore has 6.25 times the sun's luminosity.

You can also use this formula to find sizes of stars.

Example B: Suppose you found a star whose absolute magnitude is +0.8 and whose spectrum shows it has twice the sun's temperature. What is the radius of the star relative to the sun's radius? **Solution:** The star's absolute magnitude is 4 magnitudes brighter than the sun, and you recall from Reasoning with Numbers 2-1 that 4 magnitudes is approximately a factor of 2.512^4, or about 40. The star's luminosity is therefore about 40 $L_\odot$. With the luminosity and temperature, you can find the radius:

$$\frac{40}{1} = \left(\frac{R}{R_\odot}\right)^2 \left(\frac{2}{1}\right)^4$$

Solving for the radius you get:

$$\left(\frac{R}{R_\odot}\right)^2 = \frac{40}{2^4} = \frac{40}{16} = 2.5$$

So the radius is:

$$\frac{R}{R_\odot} = \sqrt{2.5} = 1.58$$

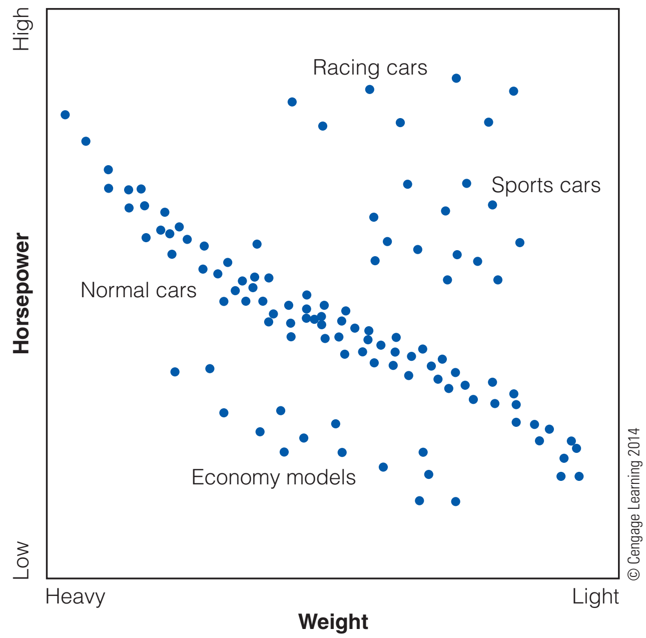
The star is 58 percent larger in radius than the sun.

*In astronomy the symbols $\odot$ and $\oplus$ refer respectively to the sun and Earth. Thus $L_\odot$ refers to the luminosity of the sun, $T_\odot$ refers to the temperature of the sun, and so on.

Karafft/Photo Researchers, Inc.

Figure 13-8

Molten lava pouring from a volcano is not as hot as a candle flame, but a lava flow has more surface area and radiates more energy than a candle flame. Approaching a lava flow without protective gear is dangerous.



■ Figure 13-9

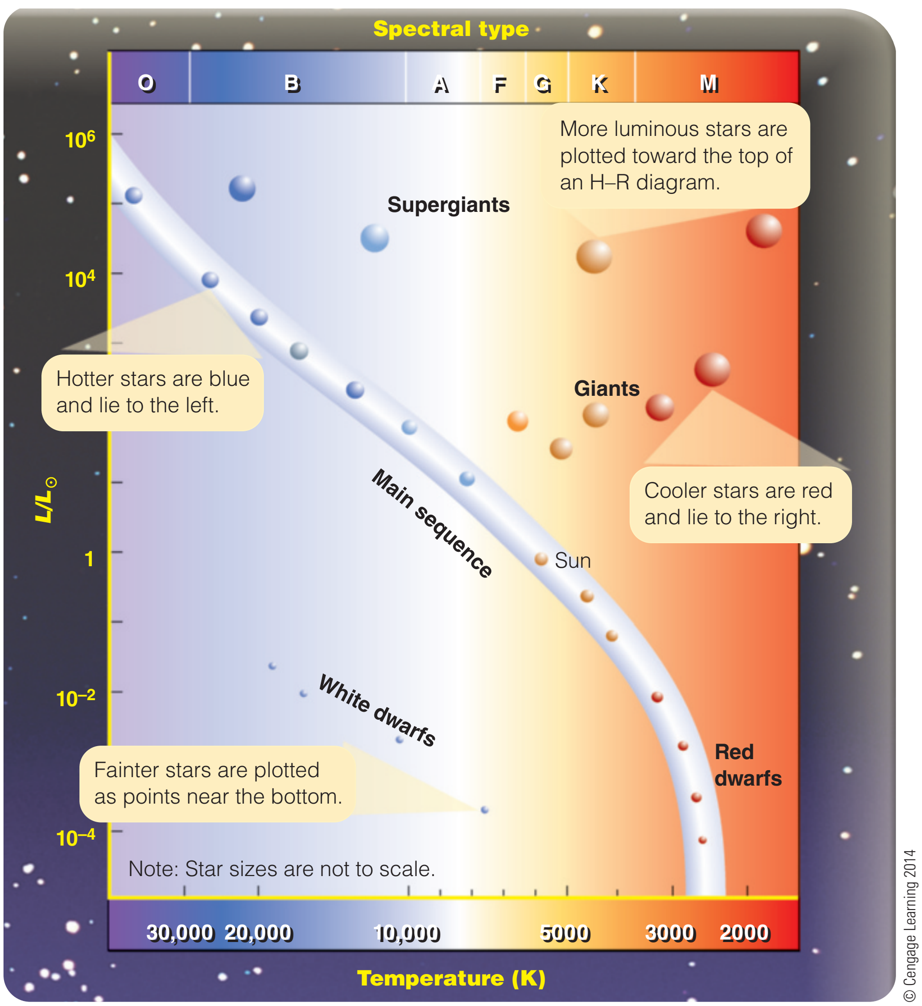
You could analyze automobiles by plotting their horsepower versus their weight and thus reveal relationships between various models. Most would lie somewhere along the main sequence of "normal" cars.

You can plot a diagram such as ■ Figure 13-9, showing horsepower versus weight for various makes of cars. In general, the more a car weighs, the more horsepower it has. Most cars fall somewhere along the sequence of ordinary cars, running from heavy, high-powered cars at the upper left to light, low-powered models at the lower right. You might call this the main sequence of cars. But some cars have much more horsepower than normal for their weight—the sport and racing models—and lie higher in the diagram. Other cars, the economy models, have less power than normal for cars of the same weight and fall lower in the diagram. Just as this diagram sorts cars into engine-power groups, so the H–R diagram sorts stars into groups according to size.

The H–R diagram is a graph with luminosity on the vertical axis and temperature on the horizontal axis. A star is represented by a point on the graph that marks its luminosity and its temperature. The H–R diagram in ■ Figure 13-10 also contains a scale of spectral type across the top. As you now know, a star's spectral type is determined by its temperature, so you can use either spectral type or temperature on the horizontal axis.

In an H–R diagram, the location of a point tells you a great deal about the star it represents. Points near the top of the diagram represent very luminous stars, and points near the bottom represent very-low-luminosity stars. Points near the right edge of the diagram represent very cool stars, and points near the left edge of the diagram represent very hot stars. Notice in the H–R diagram in Figure 13-10 how the artist has used color to represent temperature. Red stars are cool, and blue stars are hot, as you learned in Chapter 6 about Wien's law for the colors of blackbody emission.

Astronomers use H–R diagrams so often that they usually skip the words "the point that represents the star." Rather, they will say that a



■ Figure 13-10

In an H–R diagram, a star is represented by a dot at a position that shows the star's luminosity and temperature. The background color in this diagram indicates the temperature of the stars. The sun is a yellow-white G2 star. Most stars fall along the main sequence running from hot luminous stars at upper left to cool low-luminosity stars at lower right.

star is located in a certain place in the diagram. The location of a star in the H–R diagram has nothing to do with the location of the star in space. Furthermore, a star is said to "move" in the H–R diagram as it ages and its luminosity and temperature change, but such motion in the diagram has nothing to do with the star's motion in space.

Giants, Supergiants, and Dwarfs

The **main sequence** is the region of the H–R diagram running from upper left to lower right. It includes roughly 80 percent of all stars. In Figure 13-10, the main sequence is represented by a curved line with dots for stars plotted along it. As you might expect, the hot main-sequence stars are more luminous than the cool main-sequence stars.

Notice in the H–R diagram that some cool stars lie above the main sequence. Although they are cool, they are luminous, and that must mean they are larger and have more surface area than main sequence stars of the same temperature. These are called **giant** stars, and they are roughly 10 to 100 times larger than the sun. There are even **supergiant** stars at the top of the H–R diagram that are over a thousand times the sun's diameter.

At the bottom of the H–R diagram lie the economy models, stars that are very low in luminosity because they are very small. At the bottom end of the main sequence, the **red dwarfs** are not only small, they are also cool, and that gives them low luminosities. In contrast, the **white dwarfs** lie in the lower left of the H–R diagram and are lower in luminosity than you would expect, given their high temperatures. That must mean they are very small. Although some white dwarfs are among the hottest stars known, they are so small they have very little surface area from which to radiate, and that limits them to low luminosities.

The equation in Reasoning with Numbers 13-3 that relates luminosity, temperature, and radius of a star can be used to draw precise lines of constant radius across the H–R diagram, and these lines slope down and to the right across the diagram because cooler stars are fainter than hotter stars of the same size. ■ Figure 13-11 plots the luminosities and temperatures of a number of well-known stars along with lines of constant radius. For example, locate the line labeled 1 $R_\odot$ (1 solar radius) and notice that it passes through the point representing the sun. Any star whose point is located along this line has a size equal to the sun's. Next, look at the rest of the stars along the main sequence. They range from a tenth the size of the sun to about ten times as large. Even though the main sequence slopes dramatically down to the right across the diagram, most main-sequence stars are similar in size. In contrast, the white dwarfs at the lower left of the diagram are extremely small—only 1/100 the diameter of the sun, only about the size of Earth—and the giants and supergiants at the upper right are extremely large compared to the stars of the main sequence.

Notice the great range of sizes among stars. The largest stars are 100,000 times larger than the tiny white dwarfs. If the sun were a tennis ball, the white dwarfs would be grains of sand, and the largest supergiants would be as big as football fields.

Luminosity Spectral Classification

Spectra of stars contains clues not only about their temperature and composition; you can also use a star's spectrum to determine whether it is a main-sequence star, a giant, or a supergiant. The larger a star is, the less dense its atmosphere is, and that affects the widths of spectral lines.

Atoms collide often in a dense gas, their energy levels become distorted, and their spectral lines are broadened. Hydrogen Balmer lines are an example (■ Figure 13-12). In the spectrum of a main-sequence star, the Balmer lines are broad because the star's atmosphere is dense and the hydrogen atoms collide often. In the spectrum of a giant star, the lines are narrower because the giant star's atmosphere is less dense, and the hydrogen atoms collide less often. In the spectrum of a supergiant star, the Balmer lines are very narrow.

You can look at a star's spectrum and tell roughly how big it is. Size categories derived from spectra are called **luminosity classes** because the size of the star is the dominating factor in determining luminosity. Supergiants, for example, are very luminous because they are very large. The luminosity classes are represented by the Roman numerals I for supergiants through V for main-sequence stars, with supergiants further subdivided into types Ia and Ib, as follows:

Luminosity Classes

Ia	Luminous supergiant
Ib	Supergiant
II	Luminous giant
III	Giant
IV	Subgiant
V	Main sequence

This classification scheme allows you to distinguish between more luminous supergiants (Iab) such as Rigel and the regular supergiants (Ib) such as Polaris, the North Star. The star Adhara is a bright giant (II), Aldebaran is a giant (III), and Altair is a subgiant (IV). Sirius and Vega, like the sun, are main-sequence stars (V). When you describe a star, its luminosity class appears after the spectral type, as in G2 V for the sun. White dwarfs don't enter into this classification because, as you will learn later, they are remnants of stars, and their spectra are very different from any of the other types. Notice that some of our Favorite Stars are unusual; next time you look at Polaris, remind yourself that it is a supergiant.

Luminosity classification is subtle and not too accurate, but it is important in modern astronomy. As you will see in the next section, luminosity classification provides a way to

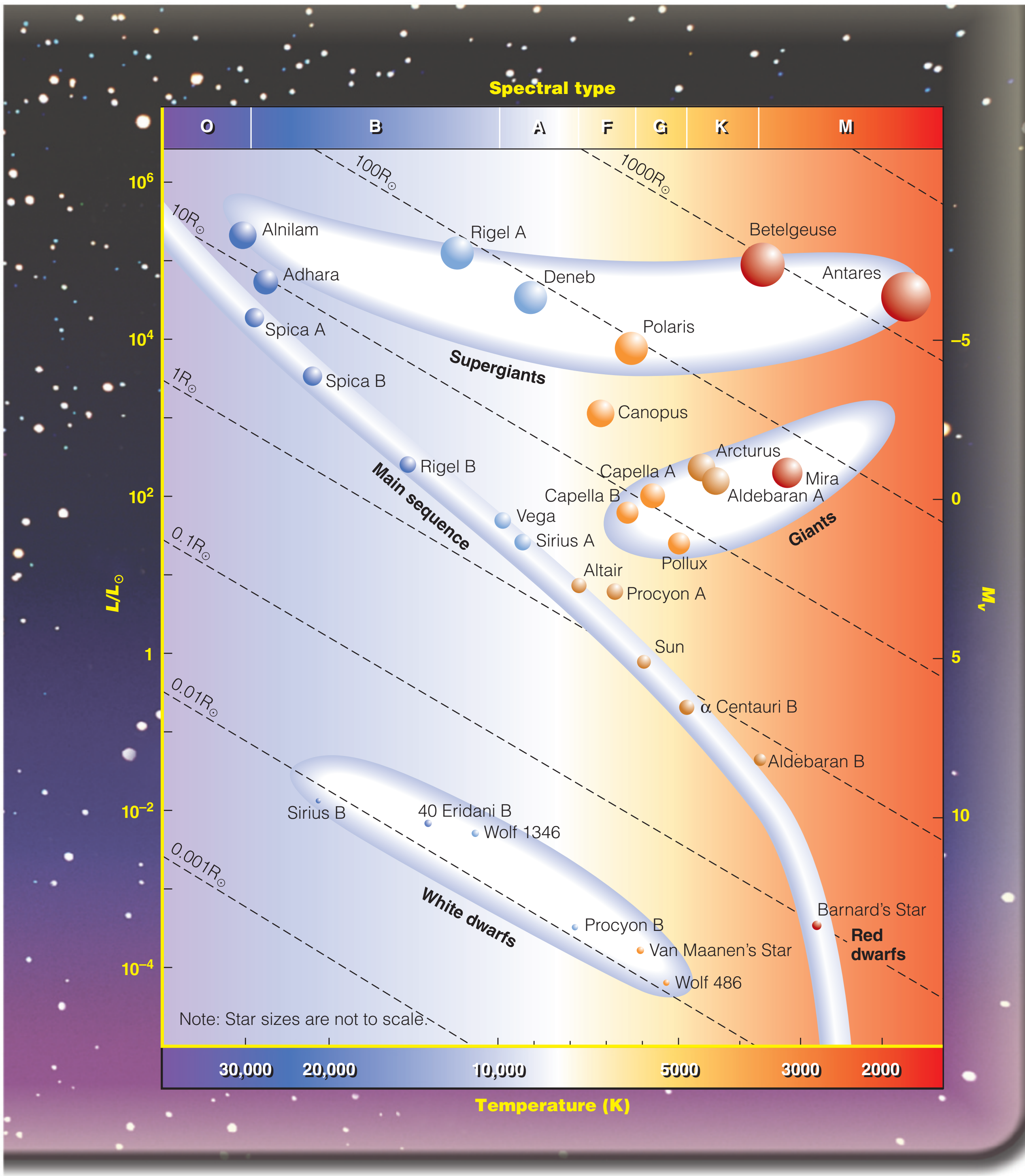


Figure 13-11

An H–R diagram showing the luminosity and temperature of many well-known stars. The dashed lines are lines of constant radius. The star sizes on this diagram are not to scale. (Individual stars that orbit each other are designated A and B, as in Spica A and Spica B.)

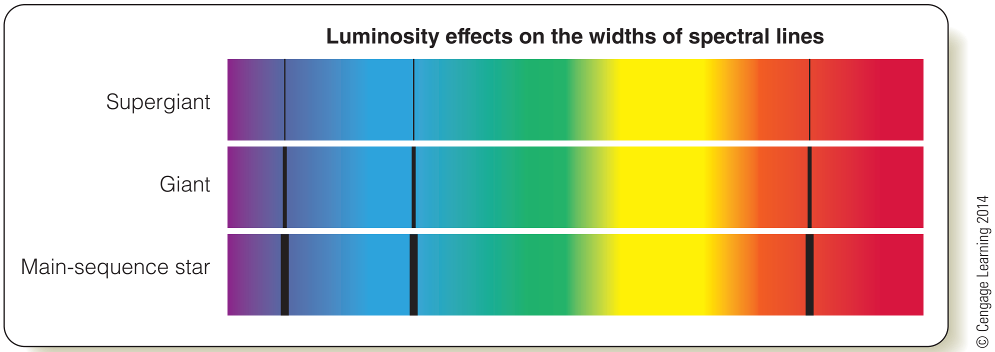


Figure 13-12

These model spectra show how the widths of spectral lines reveal a star's luminosity class. Supergiants have very narrow spectral lines, and main-sequence stars have broad lines. Certain spectral lines are more sensitive to this effect than others.

estimate the distance to stars that are too far away to have measurable parallaxes.

Spectroscopic Parallax

Astronomers can measure the stellar parallax of nearby stars, but most stars are too distant to have measurable parallaxes. These distances can be estimated from the star's spectral type, luminosity class, and apparent magnitude in a process called **spectroscopic parallax.** Spectroscopic parallax is potentially a confusing term because it does not involve an actual measurement of parallax shifts, but it is a method to tell you the distance to the star.

Spectroscopic parallax relies on the location of the star in the H–R diagram. If you record the spectrum of a star, you can determine its spectral class, and that tells you its horizontal location in the H–R diagram. You can also determine its luminosity class by looking at the widths of its spectral lines (Figure 13-12), and that allows you to estimate the star's vertical location in the diagram. Once you plot the point that represents the star in the H–R diagram, you can read off its absolute magnitude. As you learned earlier in this chapter, you can find the distance to a star by comparing its apparent and absolute magnitudes.

For example, Favorite Star Betelgeuse has an apparent magnitude of about 0.6. It is classified M2 Ib. You can find Betelgeuse plotted in Figure 13-11 at an absolute magnitude of about −6.0. The difference between Betelgeuse's apparent and absolute magnitudes, $m_V - M_V$, is therefore 0.6 minus (−6.0), or 6.6, so its distance calculated from the magnitude-distance formula (Reasoning with Numbers 13-2) is about 210 pc. A combination of parallax measurements made by the *Hipparcos* satellite and radio telescopes yields a distance of 197 pc with an uncertainty of 24 percent, so the result derived from the spectroscopic parallax method is not too bad. Obviously a real measurement of the parallax is better, but for distant stars spectroscopic parallax is often the only way to find their distance.

SCIENTIFIC ARGUMENT

What evidence can you give that giant stars really are bigger than the sun?

Scientific arguments are based on evidence, so you need to proceed step-by-step here. Stars exist that have the same spectral type as the sun but are clearly more luminous. Capella, for example, is a G star with an absolute visual magnitude of −0.5. Because it is a G star, it must have about the same temperature as the sun, but its absolute magnitude, based on its measured distance, is a bit more than five magnitudes brighter than the sun's. A magnitude difference of five corresponds to a flux ratio of 100, so Capella must be about 100 times more luminous than the sun. If it has the same surface temperature as the sun but is 100 times more luminous, then it must have a surface area 100 times greater than the sun's. Because the surface area of a sphere is proportional to the square of the radius, Capella must be ten times larger in radius. That is clear observational evidence that Capella is a giant star.

In Figure 13-11, you can see that Procyon B is a white dwarf slightly warmer than the sun but about 10,000 times less luminous. Build a scientific argument based on evidence to resolve this question. **Why do astronomers conclude that white dwarfs must be small stars?**

13-5 Star Masses—Binary Stars

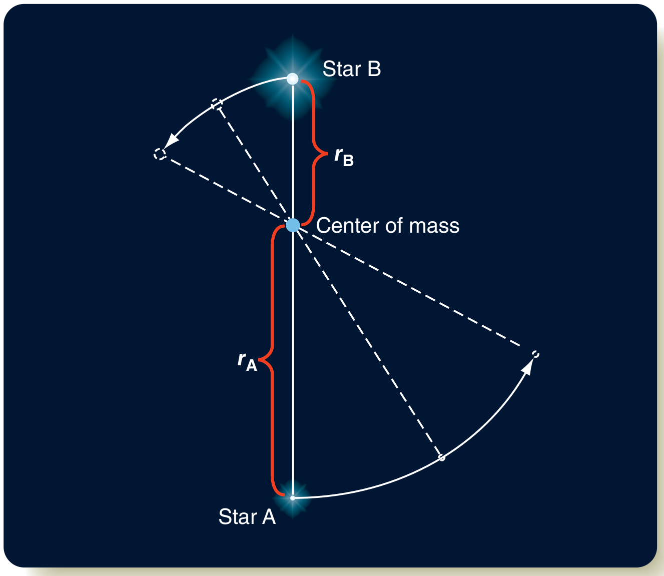
To FURTHER UNDERSTAND STARS, you need to find out how much matter they contain, that is, their masses. Gravity is the key to determining star masses. Matter produces a gravitational field, and you can figure out how much matter a star contains if you watch another object move through that star's gravitational field. To find the masses of stars, you need to study **binary stars,** pairs of stars that orbit each other.

Binary Stars in General

To measure the mass of a binary star, you first need to understand orbital motion. Chapter 4 (pages 66–67) illustrates orbits by asking you to imagine a cannonball fired from a high mountain. If Earth's gravity didn't act on the cannonball, it would follow a straight-line path and leave Earth forever. Because Earth's gravity pulls the cannonball away from its straight-line path, the cannonball follows a curved path around Earth—an orbit. When two stars orbit each other, their mutual gravitation pulls them away from straight-line paths and causes them both to follow closed orbits around a point between the stars (■ Figure 13-13).

To find the total mass of a binary star system, you must know the sizes of the orbits and the orbital period—the length of time the stars take to complete one orbit. The larger the orbits are, or the shorter the orbital period is, the stronger the stars' gravity must be to hold each other in orbit. For example, if two stars whirl rapidly around each other in small orbits, then their gravity must be very strong to prevent their flying away from each other. Such stars would have to be very massive. From the size of the orbits and the orbital period, you can figure out the masses of the stars as explained in **Reasoning with Numbers 13-4**.

Each star in a binary system moves in its own orbit around the system's center of mass, the balance point of the system. If one star is more massive than its companion, then the more massive star is closer to the center of mass and travels in a smaller orbit, while the lower-mass star whips around in a larger orbit.



■ **Figure 13-13**

As stars in a binary star system revolve around each other, the line connecting them always passes through the center of mass, and the more massive star is always closer to the center of mass.

Reasoning with Numbers | 13-4

The Masses of Binary Stars

Johannes Kepler's third law of orbital motion worked only for the planets in our solar system. When Isaac Newton realized that mass produces the gravitational attraction that causes orbits, he made that law into a general principle. Newton's version of Kepler's third law applies to any pair of objects that orbit each other. The total mass of the two objects is related to the average distance a between them and their orbital period P. If the masses are M_A and M_B, then

$$M_A + M_B = \frac{a^3}{P^2}$$

In this formula, a is expressed in AU, P in years, and the mass in solar masses.

Notice how this formula is related to Kepler's third law of planetary motion (look back to Table 4-1 on page 58). Almost all of the mass of the solar system is in the sun. If you apply this formula to any planet in our solar system, the total mass is 1 solar mass. Then the formula becomes $P^2 = a^3$, which is Kepler's third law.

This formula lets you find the masses of binary stars. You must know the distance to the binary system to be able to convert the average angular separation between the two stars into AU. Using that and their orbital period in units of years, the sum of the masses of the two stars in solar units is just a^3/P^2.

Example A: If you observe a binary system with a period of 32 years and an average separation of 16 AU, what is the total mass? **Solution:** The total mass equals $16^3/32^2$, which equals 4 solar masses.

Example B: Let's call the two stars in the previous example A and B. Suppose star A is 12 AU away from the center of mass, and star B is 4 AU away. What are the individual masses? **Solution:** The ratio of the masses must be 12:4, which equals a ratio of 3:1. What two numbers add up to 4 and have the ratio 3:1? Star B must be 3 solar masses, and star A must be 1 solar mass.

The ratio of the masses of the stars M_A/M_B equals r_B/r_A, the inverse of the ratio of the radii of the orbits. If one star has an orbit twice as large as the other star's orbit, then it must be half as massive. The period and total size of the two stars' orbits tells you the sum of their masses, and the relative sizes of the two orbits give you the ratio of their masses. That is enough to let you determine the individual masses of each star.

Figuring out the masses of stars of a binary star system has some further complications. The orbits of the two stars may be elliptical, and although the orbits lie in the same plane, that plane can be tipped at an unknown angle to your line of sight, further

How Do We Know? | 13-1

Chains of Inference

How do scientists measure something they can't detect? Sometimes scientists cannot directly observe the things they really want to study, so they must construct chains of inference that connect observable parameters to the unobservable quantities they want to know. You can't measure the mass of a star directly, so you must find a way to use what you can observe, orbital period and angular separation, to figure out step by step the parameters you need to calculate the mass.

Consider another example. Geologists can't measure the temperature and density of Earth's interior directly. There is no way to drill a hole to Earth's center and lower a thermometer or recover a sample. However, the speed of vibrations from a distant earthquake depends on the temperature and density of the rock they pass through. Geologists can't measure the speed of the vibrations deep inside Earth; but they can measure the delays in the arrival times at different locations on the surface, and that allows them to work their way back to the speed and, finally, the temperature and density.

Chains of inference can be nonmathematical. Biologists studying the migration of whales can't follow individual whales for years at a time, but they can observe them feeding and mating in different locations; take into consideration food sources, ocean currents, and water temperatures; and construct a chain of inference that leads back to the seasonal migration pattern for the whales.

This chapter contains a number of chains of inference. Almost all fields of science use chains of inference. When you can link the observable parameters step by step to the final conclusions, you gain a strong insight into the nature of science.

USGS

San Andreas fault: A chain of inference connects earthquakes to conditions inside Earth.

distorting the observed shapes of the orbits. Astronomers must find ways to correct for these distortions. In addition, astronomers analyzing binary systems must first find the distances to them so they can estimate the true size of the orbits in astronomical units. Finding the masses of binary stars requires a number of steps to get from what you can observe to what you really want to know, the masses. Constructing such sequences of steps is an important part of science (**How Do We Know? 13-1**).

Although there are many different kinds of binary stars, three types are especially useful for determining stellar masses. These are discussed separately in the next sections.

Visual Binary Systems

In a **visual binary** system, the two stars are separately visible in the telescope. Only a pair of stars with large orbits can be separated visually. If the orbits are small, the telescope cannot resolve the star images, and you see only a single point of light (look back to Chapter 5). In a visual binary system, you can see each star moving around its own orbit around the center of mass.

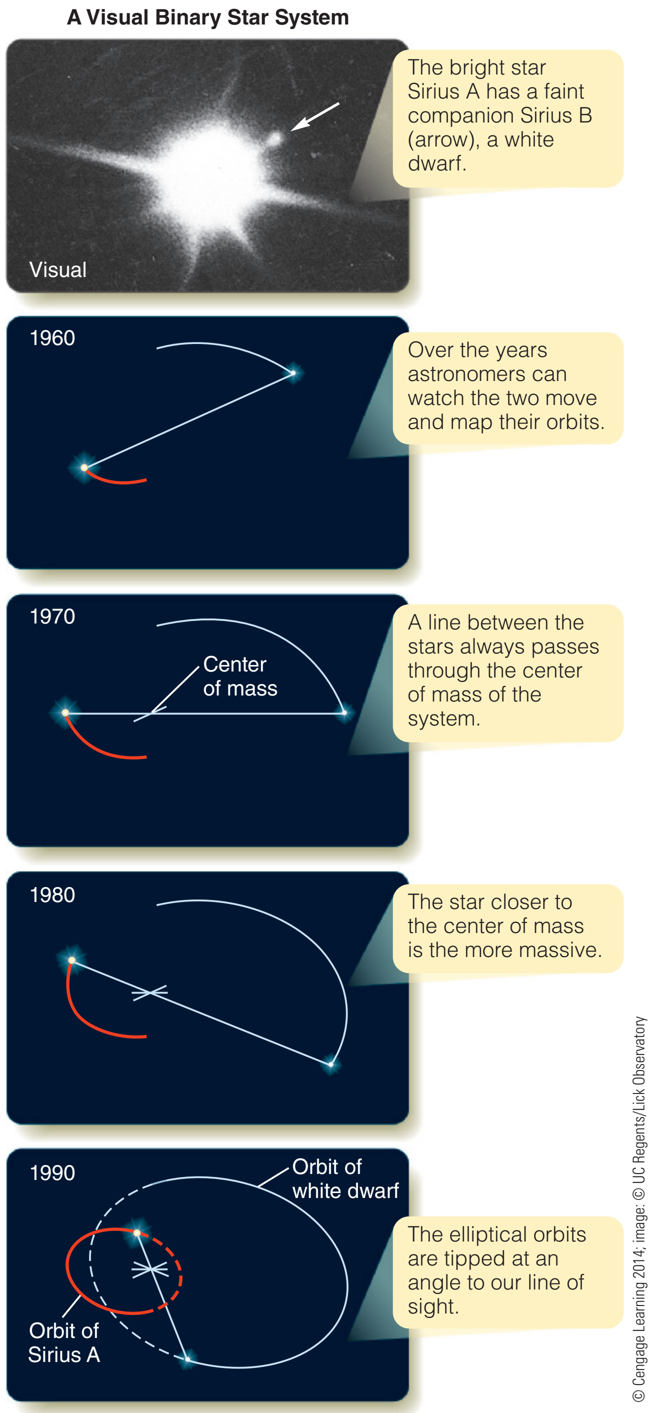
Astronomers study visual binary systems by measuring the positions of the two stars over the course of many years to map the orbits. The first frame of ■ Figure 13-14 shows a photograph of Favorite Star Sirius, which is a visual binary system made up of the bright star Sirius A and its white dwarf companion Sirius B. The photo was taken in 1960. Successive frames in Figure 13-14 show the motion of the two stars as observed since 1960 and the orbits the stars follow. The orbital period is 50 years, and astronomers have found accurate masses for both stars.

Visual binary systems are common; more than half of all stars are members of binary star systems, and many of those are visual binaries. However, many of those systems cannot be analyzed completely. For example, Favorite Star Polaris has two stellar companions. One of the stars that orbits Polaris has an orbital period estimated to be over a thousand years, so the exact size and period are not well known. The other companion is so close to Polaris that it is hardly distinguishable even with the *Hubble Space Telescope.*

Spectroscopic Binary Systems

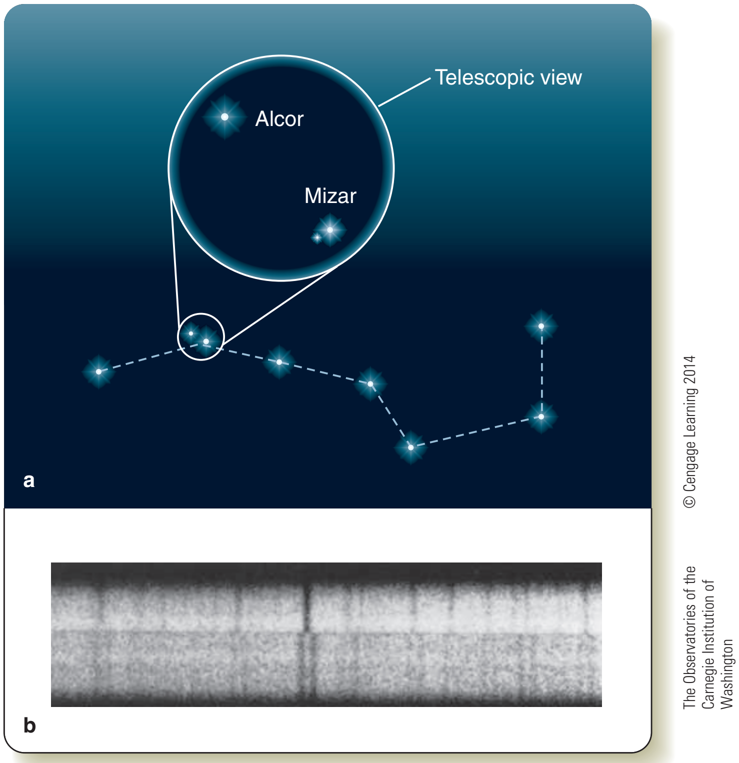
If the stars in a binary system are very close together, telescopes limited by diffraction and seeing may reveal only a single point of light. By looking at a spectrum, which is formed by light from both stars and contains spectral lines from both, astronomers can sometimes tell that there are two stars present and not one. Such a pair of stars is a **spectroscopic binary.** Most known binary systems are spectroscopic binaries. Many of the familiar stars in the sky, apparently single stars, are actually two or more stars orbiting each other (■ Figure 13-15).

Figure 13-15b shows photographs of two spectra of the star Mizar recorded on different occasions, displaying one spectral line at one time and the same spectral line split into two lines at another time. Finding pairs of spectral lines moving back and forth across each other alerts you that you are observing a spectroscopic binary system. Both Mizar and its close faint visual companion (Figure 13-15a) are now known to be spectroscopic binary systems.



■ **Figure 13-14**

The orbital motion of Sirius A and Sirius B can reveal their individual masses.



■ **Figure 13-15**

(a) At the bend of the handle of the Big Dipper lie a pair of stars, Mizar and Alcor. Through a telescope you can discover that Mizar has a fainter companion and so is a member of a visual binary system. Adaptive optics observations have discovered a faint close companion of Alcor, not pictured in this diagram. (b) Spectra of Mizar recorded at different times show that it a spectroscopic binary system rather than a single star.

To understand what's going on in Figure 13-15, look at the diagrams in ■ Figure 13-16, showing two stars orbiting each other. In the first frame, star A is approaching while star B recedes. In the cartoon spectrum below the orbit, you see a spectral line from star A Doppler shifted toward the blue end of the spectrum while the same spectral line from star B is Doppler shifted toward the red end of the spectrum (look back to Chapter 6). As the two stars continue to revolve around their orbits, they alternately approach and recede, and the changing Doppler shifts move their spectral lines to different wavelengths, apart and then back together. When the stars are in the parts of their orbits in which their velocities are across our line of sight, neither coming toward us nor going away, there are no Doppler shifts and the two lines are at the same wavelength, appearing as one line.

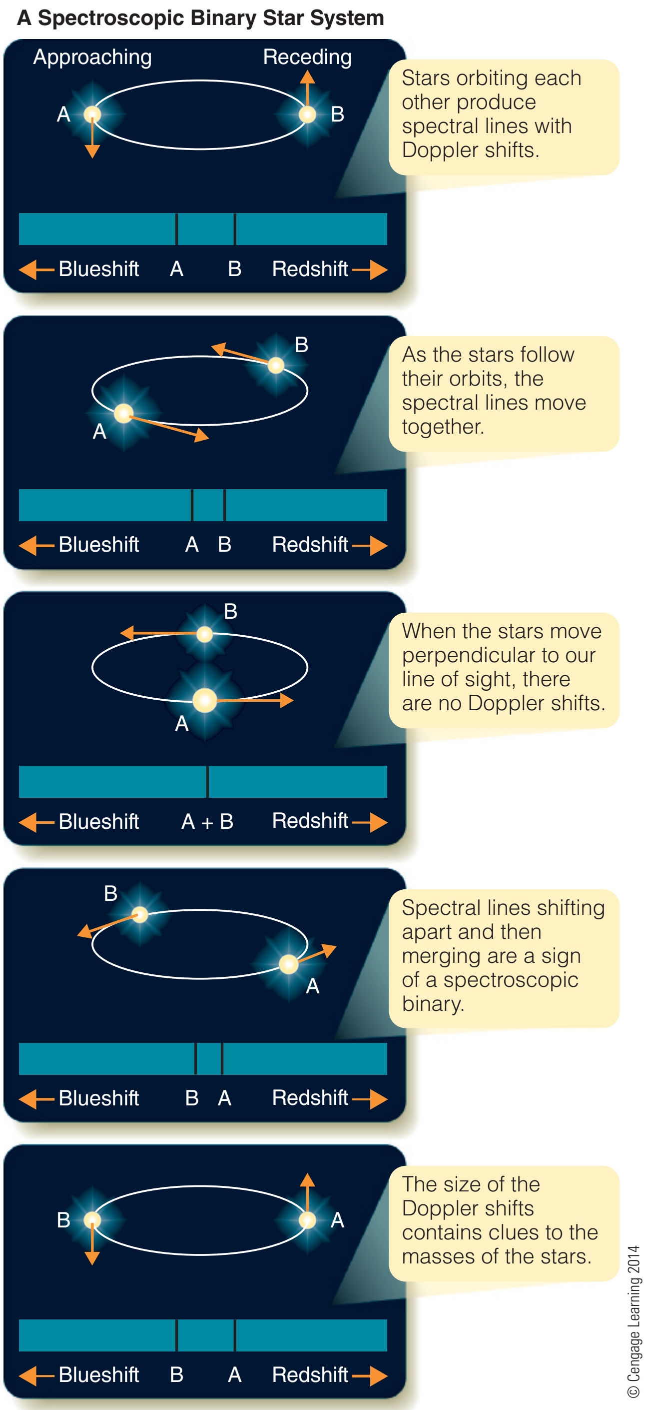
At first glance, it seems that it should be easy to find the masses of the stars in a spectroscopic binary by observing the changing Doppler shifts of the spectral lines to determine the orbital period and velocities of the stars. If you multiply velocity times orbital period, you find the circumference of the orbit. Now that you know the orbital period and the size of the orbit, you would think you would be able to calculate the mass. One important detail is missing, however. You don't know how much the orbits are inclined to your line of sight.

Astronomers can find the inclination of a visual binary system because they can see the shape of the orbits. In a spectroscopic binary system, however, the individual stars are not visible, so the orbits can't be mapped and the inclination can't be found. Recall that the Doppler effect only reveals the radial velocity, the part of the velocity directed toward or away from the observer. Because you cannot find the inclination, you cannot do a geometric correction to the radial velocities to find the true orbital velocities. Consequently, you cannot find the true masses. All you can find from a spectroscopic binary system is a lower limit to the masses.

You might wonder what happens when the orbits of a spectroscopic binary system lie exactly edge-on to Earth. The result is the most useful kind of binary system.

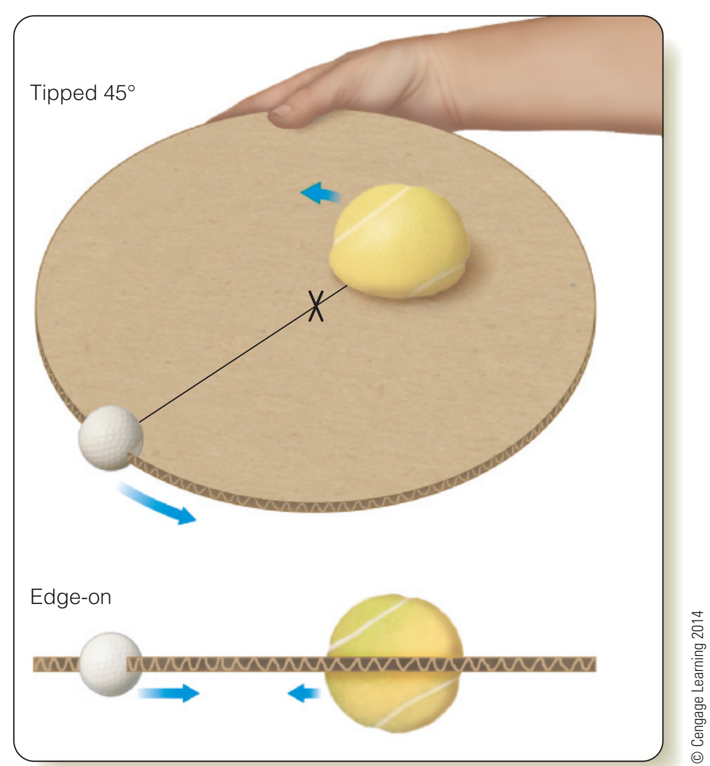
Eclipsing Binary Systems

As mentioned earlier, the orbits of the two stars in a binary system always lie in a single plane. If that plane is nearly edge-on to Earth, then the stars will appear to cross in front of each other. Imagine a model of a binary star system in which a cardboard disk represents the orbital plane and balls represent the stars, as in ■ Figure 13-17. If you see the model from the edge, then the balls that represent the stars can move in front of each other as they follow their orbits. Seen from Earth, the two stars are never visible separately, only as a single point of light. But as the stars repeatedly block each other's light from our point of view, the total brightness of the point of light periodically decreases. Such a pair of stars is called an **eclipsing binary.**



■ **Figure 13-16**

From Earth, a spectroscopic binary looks like a single point of light, but the Doppler shifts in its spectrum reveal the orbital motion of the two stars.



■ **Figure 13-17**

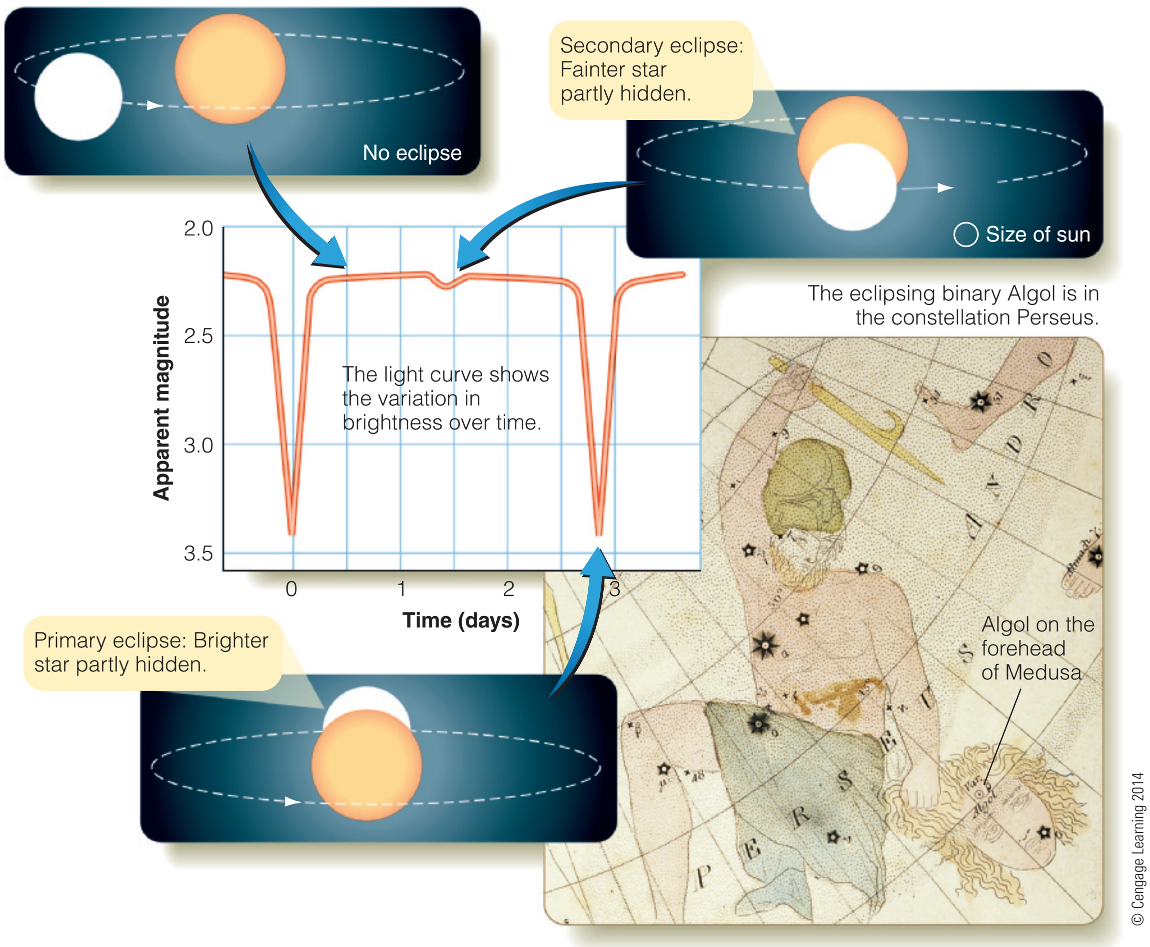
Imagine a model of a binary system with tennis and golf balls for stars and a disk of cardboard for the plane of the orbits. Only if you view the system edge on do you see the stars cross in front of each other.

Algol (Beta Persei) is one of the best-known eclipsing binaries because its eclipses are visible to the naked eye. Normally, its magnitude is about 2.15, but its brightness drops to 3.4 during eclipses that occur every 68.8 hours. Although the nature of the star as an eclipsing binary was not recognized until 1783, its periodic dimming was probably known since ancient times. *Algol* comes from the Arabic for "the demon," and it is associated in constellation mythology with the severed head of Medusa, the sight of whose serpentine locks turned mortals to stone (■ Figure 13-18). Indeed, in some accounts, the variable star Algol is the winking eye of the demon.

Figure 13-18 shows the smaller component of Algol moving in an orbit around the larger component, first eclipsing the larger star and then being eclipsed itself as it moves behind its companion. The resulting variation in the brightness of the system is shown as a graph of brightness versus time, a **light curve.** Cover the stars in Figure 13-18 with your fingers and look only at the light curve. If you saw such a light curve, you would immediately recognize that what you thought was a single star in the sky is actually an eclipsing binary.

The light curves of eclipsing binary systems can be difficult to analyze but they contain tremendous amounts of information. ■ Figure 13-19 shows the light curve of a binary system in which the stars have dark spots on their surfaces, analogous to enormous sunspots, and are so close to each other that their shapes are distorted. The light curve of this system allows crude maps of the two stars to be derived.

Once the light curve of an eclipsing binary system has been accurately observed, you can construct a chain of inference leading to the masses of the two stars. You can find the orbital period easily, and you can get spectra showing the Doppler shifts of the two stars. You can find the true orbital velocity because you don't have to untip the orbits; you know they are nearly edge-on or there would not be eclipses. Then you can find the size of the orbits and the masses of the stars.



■ **Figure 13-18**

The eclipsing binary Algol consists of a hot B star and a cooler G or K star. The eclipses are partial, meaning that neither star is completely hidden during eclipses. The orbit here is drawn as if the cooler star were stationary.

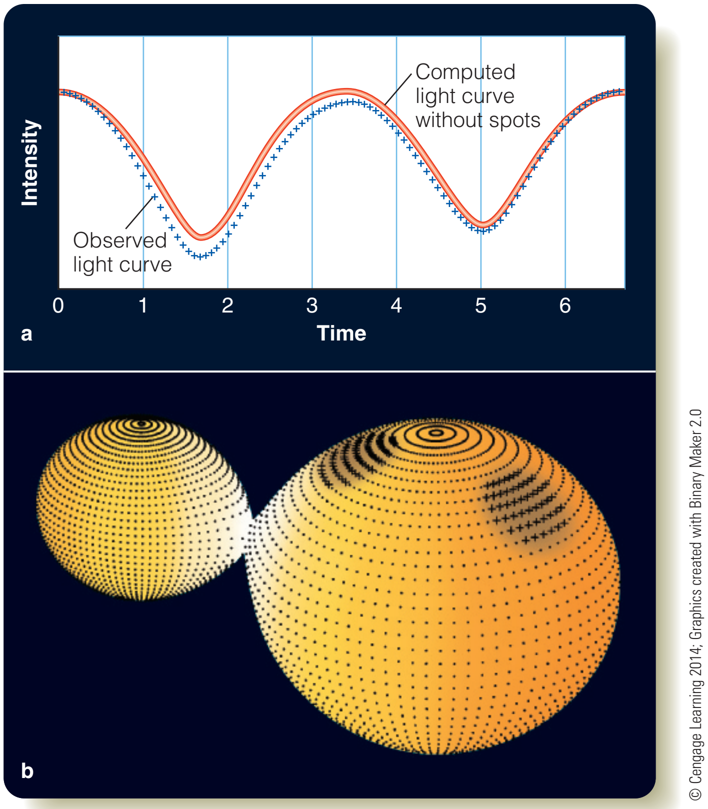


Figure 13-19

The observed light curve of the binary star VW Cephei (pronounced *sef-ee-eye*) (lower curve) shows that the two stars are so close together their gravity distorts their shapes. Slight distortions in the light curve reveal the presence of dark spots at specific places on the star's surface. The upper curve shows what the light curve would look like if there were no spots.

Earlier in this chapter you used luminosity and temperature to find the radii of stars, but eclipsing binary systems provide a way to measure the sizes of stars directly. From the light curve you can tell how long it takes for the small star to cross the large star. Multiplying this time interval by the orbital velocity of the small star gives the diameter of the larger star. You can also determine the diameter of the small star by noting how long it takes to disappear behind the edge of the large star. For example, if it takes 300 seconds for the small star to disappear while traveling 500 km/s relative to the large star, then it must be 150,000 km in diameter.

From the study of binary stars, astronomers have found that the masses of stars range from roughly 0.1 solar mass at the low end to over 100 solar masses at the high end. The most massive stars ever found in binary star systems have about 150 solar masses. A few other stars may be even more massive, but they are not members of binary systems, so astronomers must estimate their masses from models based on their luminosities, temperatures, and other characteristics.

SCIENTIFIC ARGUMENT

When you look at the light curve for an eclipsing binary system with total eclipses, how can you tell which star is hotter?

A good scientific argument is a chain of inference with no missing links. Start by assuming that the two stars in an eclipsing binary system are not the same size, so they can be labeled the larger star and the smaller star. When the smaller star moves behind the larger star, you lose the light coming from the total area of the small star. When the smaller star moves in front of the larger star, it blocks off light from the same amount of area on the larger star. In both cases, the same area (the same number of square meters) is hidden from your sight. That means the amount of light lost during an eclipse depends only on the temperature of the hidden surface because temperature is what determines how much a single square meter radiates per second. When the surface of the hotter star is hidden, the brightness will fall dramatically, but when the surface of the cooler star is hidden, the brightness will not fall as much. So you can look at the light curve and point to the deeper of the two eclipses and say, "That is where the hotter star is behind the cooler star."

Now change the argument to consider the diameters of the stars. **How could you look at the light curve of an eclipsing binary with total eclipses and find the ratio of the diameters?**

13-6 A Census of the Stars

You have learned how to find the luminosities, temperatures, sizes, and masses of stars, and now you can put all those data together (**How Do We Know? 13-2**) to paint a family portrait of the stars. As in any family portrait, both similarities and differences are important clues to the history of the family. As you begin trying to understand how stars are born and how they die, ask a simple question: What is the average star like? Answering that question is both challenging and illuminating.

Surveying the Stars

If you want to know what the average person thinks about a certain subject, you take a survey. If you want to know what the average star is like, you need to survey the stars. Such surveys reveal important relationships among the family of stars.

Not many decades ago, surveying large numbers of stars was an exhausting task, but modern computers have changed that. Specially designed telescopes controlled by computers can make millions of observations per night, and high-speed computers can compile these data into easy-to-use databases. Those surveys produce mountains of data that astronomers can "mine" while searching for relationships within the family of stars.

What could you learn about stars from a survey of the stars near the sun? Astronomers have evidence that the sun is in a typical place in the universe. Therefore, such a survey could reveal general characteristics of stars everywhere and might reveal

How Do We Know? 13-2

Basic Scientific Data

Where do large masses of scientific data come from? In one sense, science is the process by which scientists examine data and search for relationships, and it sometimes requires large amounts of data. For example, astronomers need to know the masses and luminosities of many stars before they can begin to understand the relationship between mass and luminosity.

Compiling basic data is one of the common forms of scientific work—a necessary first step toward scientific analysis and understanding. An archaeologist may spend months or even years diving to the floor of the Mediterranean Sea to study an ancient Greek shipwreck. She will carefully measure the position of every wooden timber and bronze fitting. She will photograph and recover everything from broken pottery to tools and weapons. The care with which she records data on the site pays off when she begins her analysis. For every hour the archaeologist spends recovering an object, she may spend days or weeks in her office, a library, or a museum identifying and understanding the object. Why was there a Phoenician hammer on a Greek ship? What does that reveal about the economy of ancient Greece?

Finding, identifying and understanding that ancient hammer contributes only a small bit of information, but the work of many scientists eventually builds a picture of how ancient Greeks saw their world. Solving a single binary star system to find the masses of the stars does not tell an astronomer a great deal about nature. Over the years, however, many astronomers have added their results to the growing data file on stellar masses. Scientific data accumulate and can be analyzed by later generations of scientists.

Michael A. Seeds

Collecting mineral samples can be hard work, but it is also fun.

unexpected processes governing the formation and evolution of stars. Study **The Family of Stars** on pages 282–283 and notice three important points:

1. Taking a survey is difficult because you must be sure you get an honest sample. If you don't survey enough stars, or if you don't notice some kinds of stars, you can get biased results.
2. Most stars are faint, and luminous stars are rare. The most common kinds of stars are the lower-main-sequence red dwarfs and the white dwarfs.
3. A survey reveals that what you see in the sky is deceptive. Stars near the sun are quite faint; but luminous stars, although they are rare, are easily visible even at great distances. Many of the brighter stars in the sky are highly luminous stars that you see even though they are far away.

The night sky is a beautiful carpet of stars, but they are not all the same.

Mass, Luminosity, and Density

If you survey enough stars and plot the data in an H–R diagram, you can see the patterns that hint at how stars are born, how they age, and how they die.

If you label an H–R diagram with the masses of the plotted stars, as in ■ Figure 13-20, you will discover that the main-sequence stars are ordered by mass. The most massive main-sequence stars are the hot stars, and as you run your eye along the main sequence, down and to the right in the diagram, you will find successively lower-mass stars until you reach the lowest-mass, coolest, faintest main-sequence stars.

Stars that do not lie on the main sequence are not in order on the H–R diagram according to mass. Giant and supergiant stars are a jumble of different masses, although supergiants tend to be more massive than giants. In contrast, all white dwarfs have about the same mass, somewhere in the narrow range of from 0.5 to about 1 solar mass.

Because of the systematic ordering of mass along the main sequence, the main-sequence stars follow a **mass–luminosity relation**—the more massive a star is, the more luminous it is (■ Figure 13-21) (see **Reasoning with Numbers 13-5,** page 284).

Notice how large the range in luminosity is. The observed range of stellar masses extends from about 0.08 solar mass to over 100 solar masses—a factor of roughly 1000. But the range of luminosities extends from about 10^{-6} to about 10^6 solar luminosities—a factor of a trillion (10^{12}). Clearly, a small difference in mass causes a large difference in luminosity.

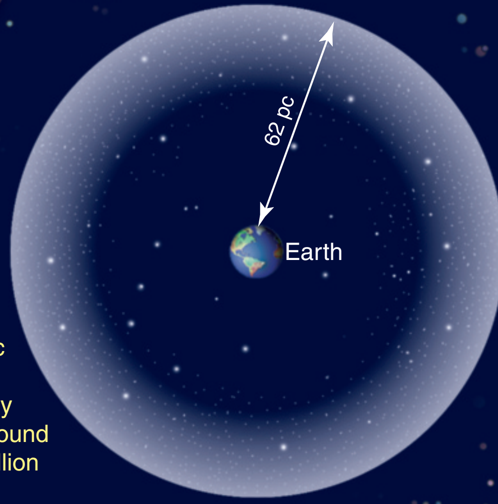
Although giants and supergiants follow different mass–luminosity relations and white dwarfs do not follow one at all, the link between mass and luminosity is critical in astronomy. In the next chapters, the mass–luminosity relation will help you understand how stars generate their energy.

The density of stars reveals another pattern in the H–R diagram. The average density of a star is its mass divided by its volume. As you will learn in the next chapter, stars are not uniform in density but are most dense at their centers and least

The Family of Stars

1 What is the most common type of star? What types of stars are rare? To answer those questions will require you to survey a large sample of stars, determining their temperature and luminosity spectral classes, apparent magnitudes, and distances. Your census of the family of stars produces some surprising demographic results.

1a You could survey the stars by observing every star within 62 pc of Earth. A sphere 62 pc in radius encloses a million cubic parsecs. Such a survey would tell you how many stars of each type are found within a volume of a million cubic parsecs.



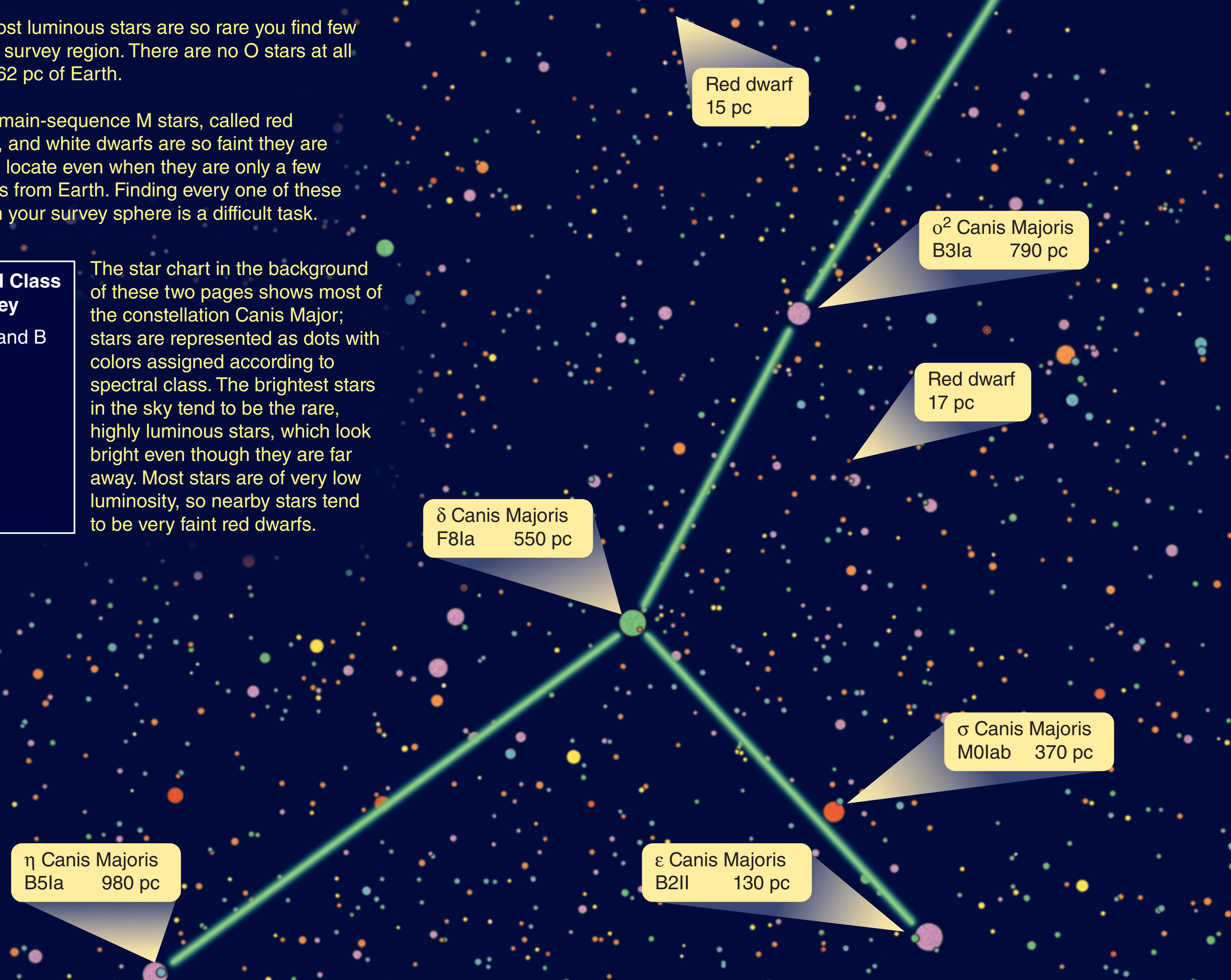
2 Your survey faces two problems.

1. The most luminous stars are so rare you find few in your survey region. There are no O stars at all within 62 pc of Earth.

2. Lower-main-sequence M stars, called red dwarfs, and white dwarfs are so faint they are hard to locate even when they are only a few parsecs from Earth. Finding every one of these stars in your survey sphere is a difficult task.

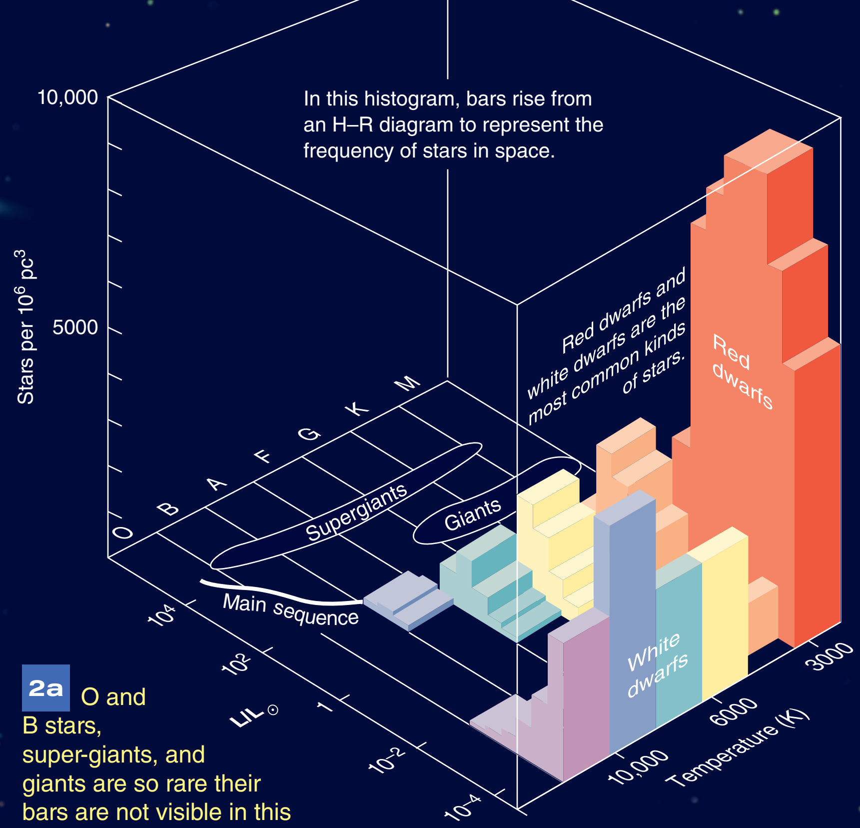
Spectral Class Color Key
- O and B
- A
- F
- G
- K
- M

The star chart in the background of these two pages shows most of the constellation Canis Major; stars are represented as dots with colors assigned according to spectral class. The brightest stars in the sky tend to be the rare, highly luminous stars, which look bright even though they are far away. Most stars are of very low luminosity, so nearby stars tend to be very faint red dwarfs.



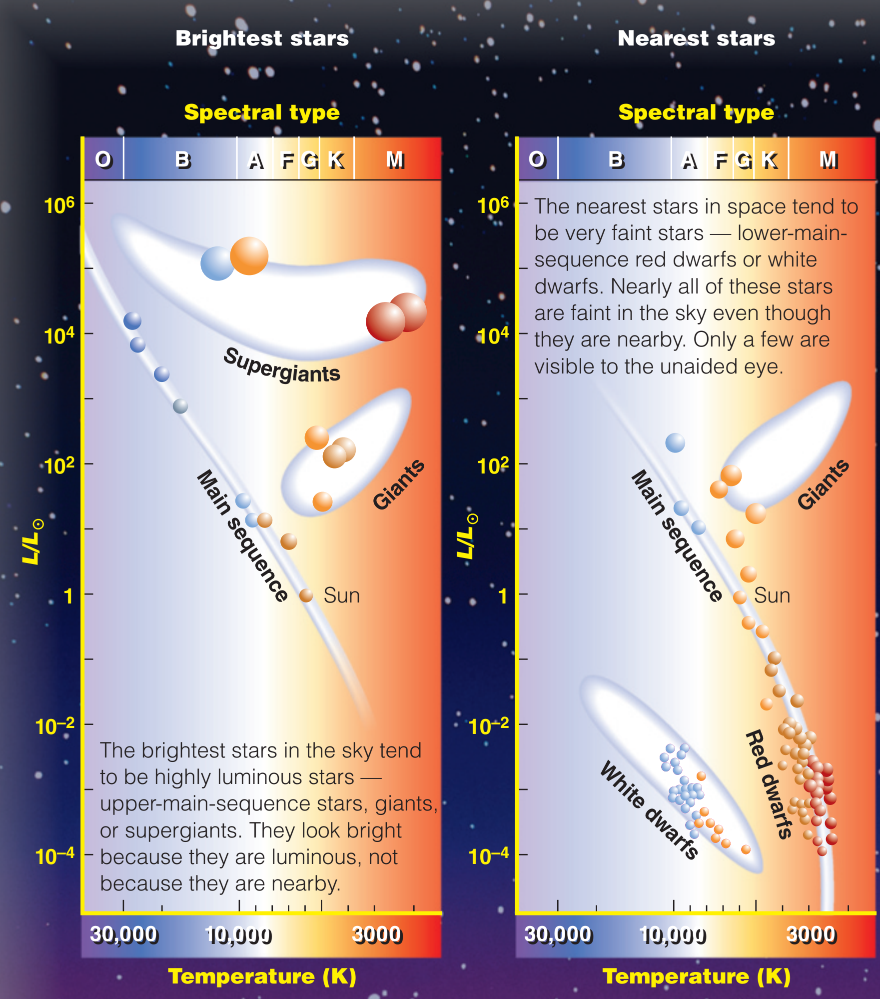
Sirius A (Alpha Canis Majoris) is the brightest star in the sky. With a spectral type of A1V, it is not a very luminous star. It looks bright because it is only 2.6 pc away.

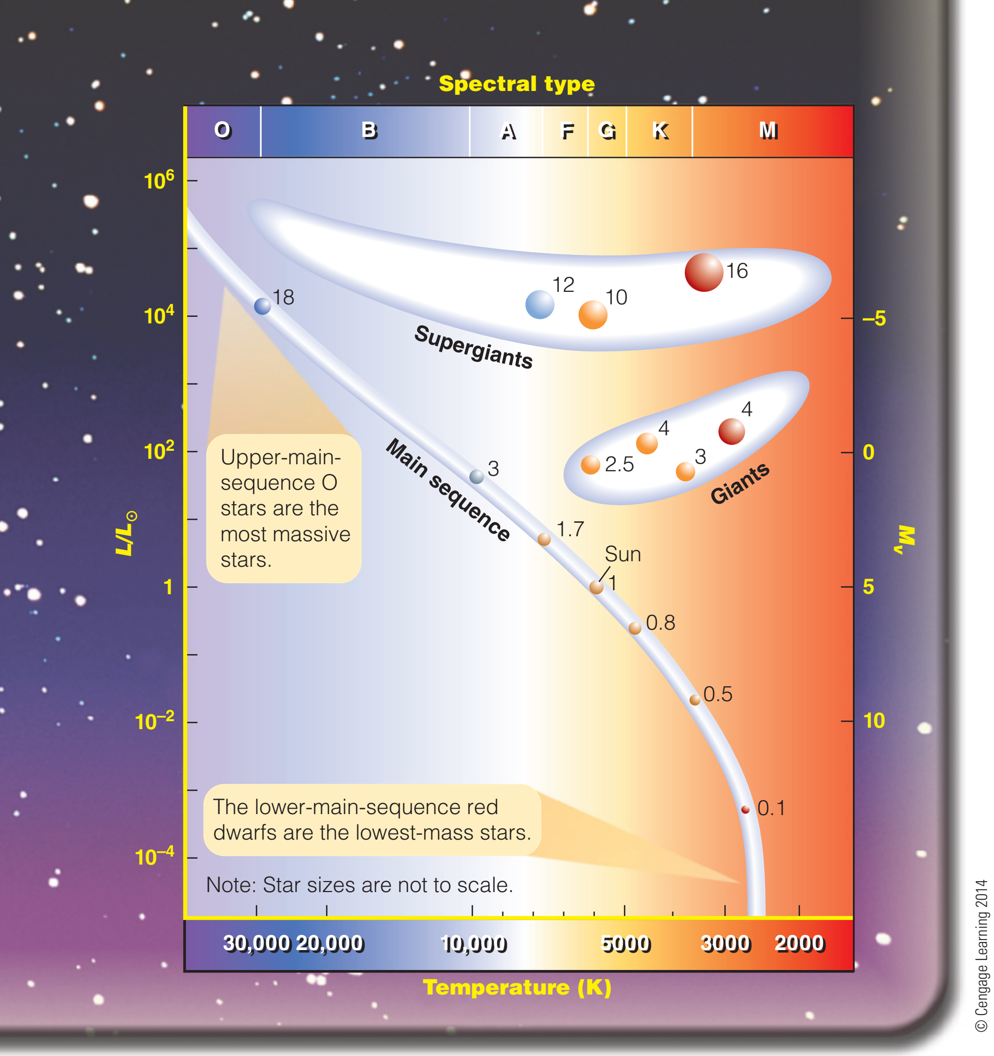
Sirius B is a white dwarf that orbits Sirius A. Although Sirius B is not very far away, it is much too faint to see with the unaided eye.



2a O and B stars, super-giants, and giants are so rare their bars are not visible in this graph.

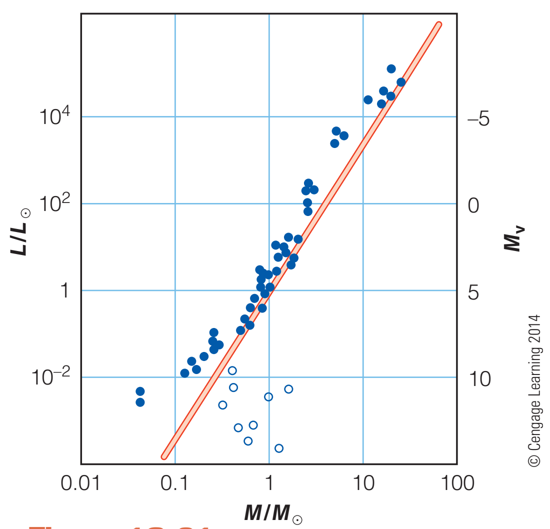
3 Luminous stars are rare but are easy to see. Most stars are very low luminosity objects. Not a single white dwarf or red dwarf is bright enough to see with the unaided eye. See H–R diagrams at right.





■ **Figure 13-20**

The masses of the plotted stars are labeled on this H–R diagram. Notice that the masses of main-sequence stars decrease from upper left to lower right in the diagram but that masses of giants and supergiants are not arranged in any ordered pattern.



■ **Figure 13-21**

The mass–luminosity relation shows that the more massive a main-sequence star is, the more luminous it is. The open circles represent white dwarfs, which do not obey the relation. The red line represents the equation $L = M^{3.5}$.

Reasoning with Numbers | 13-5

The Mass–Luminosity Relation

You can estimate the luminosity of a main-sequence star based on the star's mass using a simple equation. A main-sequence star's luminosity in terms of the sun's luminosity equals its mass in solar masses raised to the 3.5 power:

$$L = M^{3.5}$$

This is the mathematical form of the mass–luminosity relation. It is only an approximation, as shown by the red line in Figure 13-21, but it applies to most main sequence stars over a wide range of stellar masses.

Example: What is the luminosity of a main-sequence star with four times the mass of the sun? **Solution:** The star is about 128 times more luminous than the sun because:

$$L = M^{3.5} = 4^{3.5} = 128$$

dense near their surface. The center of the sun, for instance, is about 100 times as dense as water; its density near the visible surface is about 3400 times less dense than Earth's atmosphere at sea level. The sun's average density is approximately 1 g/cm^3—about the density of water, intermediate between its central and surface densities.

Main-sequence stars have average densities similar to the sun's, but giant stars have low average densities, ranging from 0.1 to 0.01 g/cm^3. The enormous supergiants have still lower densities, ranging from 0.001 to 0.000001 g/cm^3. These densities are thinner than the air you breathe, and if you could insulate yourself from the heat, you could fly an airplane through these stars. Only near their centers would you be in any danger, for there the material is very dense—millions of times denser than water.

The white dwarfs have masses about equal to the sun's but are very small, only about the size of Earth. That means the matter is compressed to enormous densities. On Earth, a teaspoonful of this material would weigh about 15 tons.

Density divides stars into three groups. Most stars are main-sequence stars with densities like the sun's. Giants and supergiants are very-low-density stars, and white dwarfs are high-density stars. You will see in later chapters that these densities reflect different stages in the evolution of stars.

SCIENTIFIC ARGUMENT

What kind of stars do you see if you look at a few of the brightest stars in the sky?

This argument shows how careful you must be to interpret simple observations. When you look at the night sky, the brightest stars are mostly giants and supergiants. Most of the bright stars in Canis Major, for instance, are supergiants. Sirius, one of your Favorite Stars, is an exception. It is the brightest star in the sky, but it is just a main-sequence star that looks bright because it is very nearby, not because it is very luminous. In general, the supergiants and giants are so luminous that they stand out and look bright, even though they are not nearby. When you look at a bright star in the sky, you are probably looking at a highly luminous star—a supergiant or a giant. You can check this argument above by consulting the tables of the brightest and nearest stars in Appendix Tables A-8 and A-9.

Now revise your argument. **What kinds of stars do you see if you look at the stars nearest to the sun?**

What Are We? Medium

We humans are medium creatures, and we experience medium things. You can see trees and flowers and small insects, but you cannot see the beauty of the microscopic world without ingenious instruments and special methods. Similarly, you can sense the grandeur of a mountain range, but larger objects, such as stars, are too big for our medium senses. You have to use your ingenuity and imagination to experience the truth of such large objects. That is what science does for us. We live between the microscopic world and the astronomical world, and science enriches our lives by revealing the parts of the universe beyond our daily experience.

Experience is fun, but it is very limited. You may enjoy a flower by admiring its color and shape and by smelling its fragrance. But the flower is more wonderful than your experience can reveal. To truly appreciate the flower you need to understand it—how complex it truly is, how it serves its plant, and how the plant came to create such a beautiful blossom.

Humans have a natural drive to understand as well as experience. You have experienced the stars in the night sky, and now you are beginning to understand them as objects ranging from hot blue O stars to cool red M dwarfs. It is natural for you to wonder why these stars are so different. As you explore that story in the following chapters, you will discover that although you have medium senses, you can understand the stars.

Study and Review

Summary

- Your goal in this chapter was to characterize the stars by finding their luminosities, temperatures, diameters, and masses. Before you could begin, you needed to find their distances. Only by first knowing the distance to a star can you find its intrinsic properties such as luminosity and diameter.
- Astronomers can measure the distance to nearer stars by observing their **stellar parallaxes *(p)* (p. 261).** The most distant stars are so far away that their parallaxes are immeasurably small. Space telescopes above Earth's atmosphere have measured the parallaxes of huge numbers of stars.
- Stellar distances are commonly expressed in **parsecs (pc) (p. 261).** One parsec is 2.06×10^5 AU—the distance to an imaginary star whose parallax is 1 arc second. One parsec equals 3.26 light-years.
- The amount of light received from a star, the light **flux (p. 262),** is related to its distance by the inverse square law. Once you know the distance to a star, you can use the **magnitude-distance formula (p. 263)** to find its **intrinsic brightness (p. 262)** expressed as its **absolute visual magnitude (M_v) (p. 262)**—the apparent magnitude the star would have if it were 10 pc away.
- To find the energy output of a star, astronomers must correct for the light at wavelengths that are not visible to convert absolute visual magnitude into **luminosity *(L)* (p. 264),** the total energy radiated by the star in one second.
- The strength of spectral lines depends on the temperature of the star. For example, in cool stars, the Balmer lines are weak because atoms are not excited out of the ground state. In hot stars, the Balmer lines are weak because atoms are excited to higher orbits or are ionized. Only at medium temperatures, around 10,000 K, are the Balmer lines strong.
- A star's temperature **spectral class** (or **type**) **(p. 266)** is determined by which absorption lines are visible in its spectrum. The **spectral sequence (p. 266),** OBAFGKM, is important because it is a temperature sequence. By classifying a star, the astronomer learns the temperature of the star's surface.
- Long after the spectral sequence was created, astronomers found the **L dwarfs (p. 266), T dwarfs (p. 266**), and **Y dwarfs (p. 266)** with temperatures even cooler than the M stars. These are examples of **brown dwarfs (p. 266),** objects smaller than stars but larger than planets.
- The **Hertzsprung–Russell (H–R) diagram (p. 269)** is a plot of luminosity versus surface temperature. It is an important diagram in astronomy because it sorts the stars into categories by size.
- Roughly 90 percent of normal stars, including the sun, fall on the **main sequence (p. 272),** with the more massive stars being hotter and more luminous. The **giants (p. 272)** and **supergiants (p. 272),** however, are much larger and lie above the main sequence.
- **Red dwarfs (p. 272)** lie at the bottom end of the main sequence. Some of the **white dwarfs (p. 272)** are very hot stars, but they fall below the main sequence because they are so small.
- Observations of a few stars made with interferometers confirm the sizes implied by the H–R diagram.
- The large sizes of the giants and supergiants mean their atmospheres have low densities and their spectra have sharper spectral lines than the spectra of main-sequence stars. In fact, it is possible to assign stars to **luminosity classes (p. 272)** by the widths of their spectral lines. Class V stars are main-sequence stars. Giant stars, class III, have sharper lines, and supergiants, class I, have extremely sharp spectral lines.
- Astronomers can use the locations of the luminosity classes in the H–R diagram to estimate the distances to stars in a technique called **spectroscopic parallax (p. 274).**
- The only direct way to find the masses of stars is by studying **binary stars (p. 274),** in which two stars orbit their common center of mass. Astronomers find the masses of the stars by observing the period and sizes of their orbits. In a **visual binary (p. 276),** both stars are visible and the orbits can be mapped, but in a **spectroscopic binary (p. 276)** the stars are so close together they look like a single point of light, and the orbits can't be observed directly. Because the inclination of the orbit can't be found, spectroscopic binaries yield only a lower limit to the masses of the stars.
- In an **eclipsing binary (p. 278),** the orbits are edge-on, and the stars cross in front of each other. The resulting brightness changes in the **light curve (p. 279)** can reveal the diameters of the stars as well as their masses.
- A survey in the neighborhood of the sun shows that lower-main-sequence stars, less luminous than the sun, are the most common type. Giants and supergiants are rare, but white dwarfs are fairly common, although they are faint and hard to find.
- The **mass–luminosity relation (p. 281)** expresses the fact that the more massive a main sequence star is, the more luminous it is. The most massive main sequence stars are type O at the upper left of the H–R diagram, and the least massive are type M at the lower right of the diagram. Giants and supergiants stars do not follow a simple mass–luminosity relation, and neither do white dwarfs.
- Given the mass and diameter of a star, can you can find its average density. Main sequence stars have about the same density as the sun, but giants and supergiants are very-low-density stars. Some are much thinner than air. The white dwarfs, lying below the main sequence, are tremendously dense.

Study and Review

Review Questions

1. Why are Earth-based parallax measurements limited to the nearest stars?
2. Why was the *Hipparcos* satellite able to make more accurate parallax measurements than ground-based telescopes?
3. What do the words *absolute* and *visual* mean in the definition of absolute visual magnitude?
4. What does luminosity measure that is different from what absolute visual magnitude measures?
5. Why are Balmer lines strong in the spectra of medium-temperature stars and weak in the spectra of hot and cool stars?
6. Why are titanium oxide features visible in the spectra of only the coolest stars?
7. Explain the interrelationships among Table 13-1, Figure 13-4c, Figure 13-5, and Figure 13-6.
8. Why does the luminosity of a star depend on both its radius and its temperature?
9. How can you be sure that giant stars really are larger than main-sequence stars?
10. What evidence shows that white dwarfs must be very small?
11. What observations would you make to classify a star according to its luminosity? Why does that method work?
12. Why does the orbital period of a binary star depend on its mass?
13. What observations would you make to study an eclipsing binary star, and what would those measurements tell you about the component stars?
14. Why don't astronomers know the inclination of a spectroscopic binary? How do they know the inclination of an eclipsing binary?
15. How do the masses of stars along the main sequence illustrate the mass–luminosity relation?
16. Why is it difficult to find out how common the most luminous stars are? The least luminous stars?
17. What is the most common type of star?
18. If you look only at the brightest stars in the night sky, what type of star are you likely to be observing? Why?
19. **How Do We Know?** What is the missing link in the chain of inference leading from observations of spectroscopic binaries to the masses of stars?
20. **How Do We Know?** In what way does accumulation of large amounts of basic scientific data help later scientists?

Discussion Questions

1. If someone asked you to compile a list of the nearest stars to the sun based on your own observations, how would you select your sample, what measurements would you make, and how would you analyze the measurements to detect nearby stars?
2. Can you think of classification systems used to simplify what would otherwise be complex measurements? Consider foods, movies, cars, grades, and clothes.
3. The sun is sometimes described as an average star. Is that true? What is the average star really like?

Problems

1. If a star has a parallax of 0.050 arc second, what is its distance in pc? In ly? In AU?
2. If a star has a parallax of 0.016 arc second and an apparent magnitude of 6, how far away is it, and what is its absolute magnitude?
3. Complete the following table of apparent visual magnitudes, absolute visual magnitudes, distances, and parallaxes:

m_V	M_V	d (pc)	p (arc seconds)
____	7	10	____
11	____	1000	____
____	−2	____	0.025
4	____	____	0.040

4. Determine the temperatures of the following stars based on their spectra. Use Figure 13-4c.
   a. medium-strength Balmer lines, strong helium lines
   b. medium-strength Balmer lines, weak ionized-calcium lines
   c. strong TiO bands
   d. very weak Balmer lines, strong ionized-calcium lines
5. To which spectral classes do the stars in Problem 4 belong?
6. If a main-sequence star has a luminosity of 400 $L_\odot$, what is its spectral type? (*Hint:* See Figure 13-10 or 13-11.)
7. If a star has an apparent magnitude equal to its absolute magnitude, how far away is it in parsecs? In light-years?
8. An O8 V star has an apparent magnitude of +1. Use the method of spectroscopic parallax to estimate the distance to the star. (*Hints:* Refer to one of the H–R diagrams in the chapter; use the magnitude-distance formula in Reasoning with Numbers 13-2.)
9. At the position of Earth, the total flux of sunlight at all wavelengths, called the solar constant, is 1366 watts per m². Use that to find the total luminosity of the sun. Make your calculations in two steps. First, use $4\pi R^2$ to calculate the surface area in m² of a sphere surrounding the sun with a radius of 1 AU. Second, multiply by the solar constant to find the total solar energy passing through the sphere in 1 second. (This assumes that the sun's luminosity is emitted equally in all directions.) The result will be the luminosity of the sun. Compare your result with that in Celestial Profile 1, Chapter 7.
10. In the following table, which star is brightest in apparent magnitude? Most luminous in absolute magnitude? Largest? Farthest away?

Star	Spectral Type	m_V
a	G2 V	5
b	B1 V	8
c	G2 Ib	10
d	M5 III	19
e	White dwarf	15

11. What is the total mass of a visual binary system if the average separation of the stars is 8 AU and their orbital period is 20 years?
12. If an eclipsing binary has a period of 32 days, and the two components are in circular orbits around their center of mass with speeds respectively of 154 km/s and 77 km/s, what are the circumferences of the two orbits? The separation between the two stars? The total mass of the system? The mass of each star?
13. What is the luminosity of a 4-solar-mass main-sequence star? Of a 8-solar-mass main-sequence star? Can you easily determine the luminosity of a 4-solar-mass red giant star? Why or why not?

Learning to Look

1. Look at Figure 13-8. Why is the lava nearest the source brighter and yellower than the lava that is farther away?
2. If all of the stars in the photo here are members of the same star cluster, then they all have about the same distance. Then why are three of the brightest much redder than the rest? What kind of star are they?


NASA

Great Debates

1. *HR Diagram.* The Hertzsprung–Russell (H–R) diagram is a plot of a star's luminosity or absolute magnitude against the star's temperature or spectral classification, respectively. Danish astronomer Ejnar Hertzsprung, in 1908, and American astronomer Henry Russell, in 1913, independently created the diagram. The diagram can be used to interpret astrophysical data, including a star's life cycle. The Nobel Prize in physics is given once per year and was first awarded in the early 1900s. Neither Hertzpring nor Russell received the Nobel Prize for their work, and awards are not given posthumously. Should these scientists have received the Nobel Prize in physics for their work? Should the posthumous requirement be lifted?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Use information you find in your research on those who have received the Nobel Prize and how the selection process works.
   c. Cite your sources.
2. *Solar Twin?* More than half of stars are in binaries, but our own solar system is not. Could the sun have a brown dwarf star in a wide orbit as its gravitational companion, and we have not discovered it yet?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Research previously suggested "solar analog candidates" to support your claim.
   c. Cite your sources.
3. *Spectral Classes L, T, Y.* Should spectral classes L, T, and/or Y be added to the current O, B, A, F, G, K, M spectral classes? The L class (1300–2000 K) has some stellar and some substellar (that is, brown dwarf) objects showing metal hydrides and alkali metals in their spectra. The T class (~700–1300 K) contains cool brown dwarfs with peak emission in the infrared and showing prominent methane in their spectra. The Y class (<600 K) is suggested to be ultracool brown dwarfs, although only six objects are currently known, and therefore this spectral class is not yet well defined. Most astronomy texts do not include these additional spectral classes. Should they? If so, what new mnemonic device do you propose for remembering the spectral classes?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.
4. *Luminosity Classes.* Luminosity classes tell us not only how bright a star is but also roughly how luminous a star is. The Yerkes spectral classification system also adds luminosity class 0 for hypergiants (i.e., the largest, most luminous and massive supergiants), VI for subdwarfs (i.e., stars with lower metal abundances, dimmer, and smaller than main sequence stars but larger and brighter than white dwarfs), and VII for white dwarfs. Is the Yerkes classification a hypothesis, theory, model, or law? Most astronomy texts do not include the Yerkes classification. Should undergraduate students in astronomy learn about the Yerkes classification?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources. Word count is 150 words or fewer.
5. *The Cannon Spectral Classification?* The Yerkes spectral classification system is so named because the authors were from Yerkes Observatory. Annie J. Cannon and Antonia C. Maury rearranged the old alphanumerical system of letters to the current temperature-based classification (see previous problem). Cannon subdivided each letter into the number 0–9 sequence with smaller numbers indicating hotter subclassification. The result was published in the *Henry Draper Catalogue* and the *Henry Draper Extension* in 1918. Back then, women were not honored for their achievements. Because Cannon led the classification, should the luminosity classification given in your text be labeled in honor of Cannon and called the Cannon Spectral Classification?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Apparent Brightness and Distance
- Binary Star Center of Mass
- Double-Line Spectroscopic Binary

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 16: Astronomical Distance Scales

Earth can't be moving! For one thing, you would feel it move, right? Another problem with the moving-Earth idea is that, if Earth goes around the sun, you should observe the stars apparently moving back and forth as your point of view shifts. In other words, if Earth circles the sun once per year, stars should have an annual parallax—their apparent positions should wiggle back and forth in a yearly cycle. But, because stars do not show any such parallax, then QED, Earth is not revolving around the sun.

This imagined argument is one that was given, in some form, to the Greek astronomer and mathematician Aristarchus (*air-is-STAR-kus*). He was the first person known to have articulated a complete heliocentric theory of the solar system, 1800 years before Copernicus. Aristarchus proposed that Earth rotates once per day and revolves around the sun once per year, just another planet, not the universe's central orb. Aristarchus's colleagues were capable of using the logic of the scientific method even though science would not be born in its modern form for almost two millennia. They could understand that Aristarchus was making a hypothesis, they could work out what the observable consequences of that hypothesis should be, and then they could note that those logical consequences were obviously not correct, so the hypothesis should be rejected. Back to the drawing board, Aristarchus.

Copernicus's heliocentric model drew the same fire: no stellar parallax, no Earth orbital motion. The answer is that distances of stars are so great that their annual parallaxes are undetectable without a telescope. As you learned in this chapter, the nearest star system is more than 4 light years away. Therefore, the largest stellar parallax, observed with a 1 AU baseline, is less than 1 arcsecond—smaller than the angular size of a dime a mile away. All other stars have even smaller parallaxes. It was almost 300 years after Copernicus published his book *De Revolutionibus* before technology improved to the point that stellar parallaxes were successfully measured. You might already have studied how the distance to the sun—the AU—was measured first using parallax measurements with the diameter of Earth as the baseline. Notice that measurements of greater distances are built on that foundation; the size of the solar system was measured in terms of the size of Earth, then distances to stars measured in terms of the size of the solar system.

Section 2 of Virtual Astronomy Lab 16, "Astronomical Distance Scales," reinforces what you learned in this chapter about parallax with an exercise to measure stellar parallaxes with Earth's orbit as the baseline, and then calculate the corresponding stellar distances in parsecs and light years. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 11: Spectral Sequence

The H–R diagram is the most important graph in astronomy. Even astronomers who study distant galaxies would agree that their research stands on the foundation of an understanding of stars. The H–R diagram is a graph of absolute magnitude versus spectral class, and it reveals some of the deepest secrets of the stars.

To make an H–R diagram, you need measurements. Determining a star's spectral type from a spectrum takes time, so astronomers often use the color of a star as a substitute. Spectral type is related to color; M stars are red and O stars are blue. Astronomers use a color index as a numerical measurement of a star's color.

One way in which astronomers use the H–R diagram is to find the distance to a star cluster. To do that, they measure the apparent magnitudes and colors of stars in the cluster and then compare those measurements with an H–R diagram to find the absolute magnitude of stars with those colors. Then they can calculate the distance modulus, the difference between apparent and absolute magnitudes, for each star. The average distance modulus of the stars in the cluster, converted to distance, is the best estimate for the distance to the cluster.

Section 1 of Virtual Astronomy Lab 11, "Spectral Sequence," helps you practice using absolute magnitudes to sort stars by luminosity and color. To do this exercise, you need to be familiar with absolute magnitude and apparent magnitude from your reading of this chapter. Note that in this exercise, unfortunately and confusingly, the phrase "most apparent brightness" refers to the star that looks the brightest to your unaided eye, with the magnitude of lowest numerical value. Section 2 allows you to calculate the distance to two star clusters by comparing the apparent magnitudes of their stars to an H-R diagram. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 12: Binary Stars

If you like puzzles, you'll like binary stars. Each system is a unique and challenging puzzle, but if you solve it you learn the masses of the stars. In fact, analyzing binary star systems is the *only* way to determine, rather then guess, the masses of stars. That's important because astronomers need to know the masses of lots of stars before they can begin to understand the internal structures and life stories of stars.

As you learned in this chapter, the two stars in a binary system orbit each other, or, to be more accurate, they orbit their common center of mass. Because like all other objects in the universe the stars must obey Newton's laws of motion and gravity, you can use Newton's version of Kepler's third law to analyze their orbits. You need to know the sizes of the orbits and the orbital period; that tells you the total mass of the two stars. (By the way, you must first know the distance to the system to know the true sizes of the orbits.) The ratio of the average distances of the two stars from the center of mass gives you the reciprocal of the ratio of the two masses. If you know the sum and the ratio of the two masses, you have enough information to figure out their individual values.

Section 1 of Virtual Astronomy Lab 12, "Binary Stars," reinforces what you have already learned from the textbook about such systems. Section 2 lets you work step-by-step, as astronomers do, observing the sizes of the orbits and the orbital period of a visual binary system to derive the masses of the stars. Section 3 advances your skills to analyze Algol, a famous eclipsing binary in which the eclipses are partial but detectable even with the unaided eye. Notice that gravity distorts the shapes of the stars from perfect spheres. These are complications, but they just make the puzzle more interesting. Section 4 let you practice on the most difficult type of binary star puzzle, spectroscopic binary systems. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

14

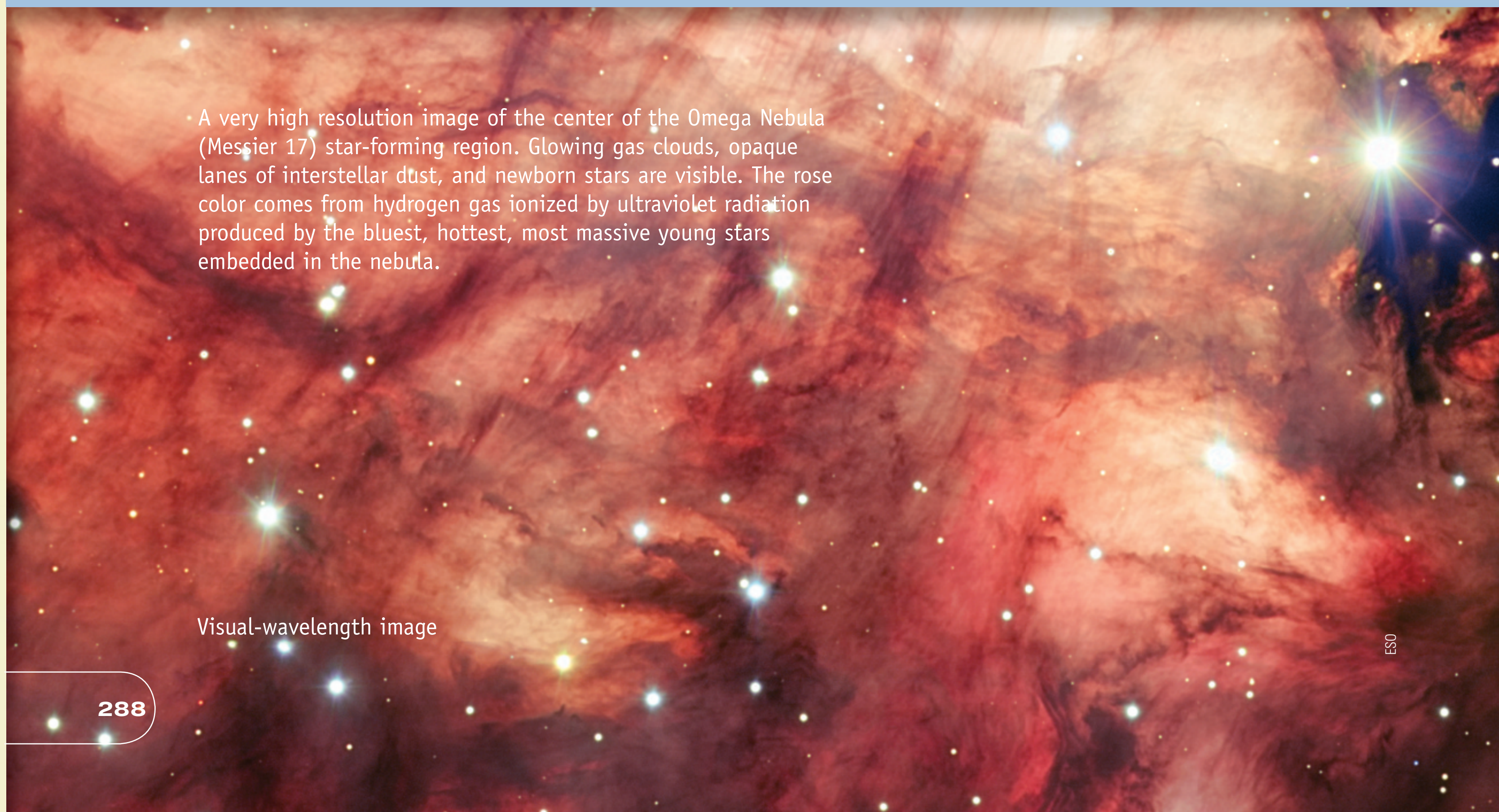
The Formation and Structure of Stars

Guidepost

In the previous chapter you discovered the wide range of differences within the family of stars. In this chapter you will combine observations and hypotheses to understand how nature makes stars, and learn the answers to five important questions:

- **How do astronomers study the gas and dust between the stars, called the interstellar medium?**
- **How do stars form from the interstellar medium?**
- **How do stars maintain their stability?**
- **How do stars make energy?**
- **How do the luminosities and lifetimes of stars depend on their masses?**

Perhaps more important than these factual questions is the key question you should ask in any scientific context: What's the evidence? Astronomers have evidence about how stars are born, how they remain stable, and how they make their luminosity. Testing hypotheses against evidence is the basic skill needed by all scientists, and you will find opportunities to use it repeatedly in the chapters that follow.



A very high resolution image of the center of the Omega Nebula (Messier 17) star-forming region. Glowing gas clouds, opaque lanes of interstellar dust, and newborn stars are visible. The rose color comes from hydrogen gas ionized by ultraviolet radiation produced by the bluest, hottest, most massive young stars embedded in the nebula.

Visual-wavelength image

ESO

Jim he allowed [the stars] was made, but I allowed they happened. Jim said the moon could'a laid them; well, that looked kind of reasonable, so I didn't say nothing against it, because I've seen a frog lay most as many, so of course it could be done.

MARK TWAIN,
THE ADVENTURES OF HUCKLEBERRY FINN

THE STARS ARE not eternal. When you look at the sky, you see hundreds of points of light, and, amazingly enough, each one, like the sun, is a tremendous nuclear fusion reactor held together by its own gravity. The stars you see tonight are the same stars your parents, grandparents, and great-grandparents saw. Stars don't change perceptibly in a person's lifetime, or even during all of human history, but they do not last forever. Stars are "born," and stars "die." This chapter begins that story.

How can you know what the life cycles and internal processes of stars are, given that you won't live long enough to see them evolve, and you can't see inside them? The answer lies in the methods of science. By constructing theories that describe how nature works and then testing those theories against evidence from observations, you can unravel some of nature's greatest secrets.

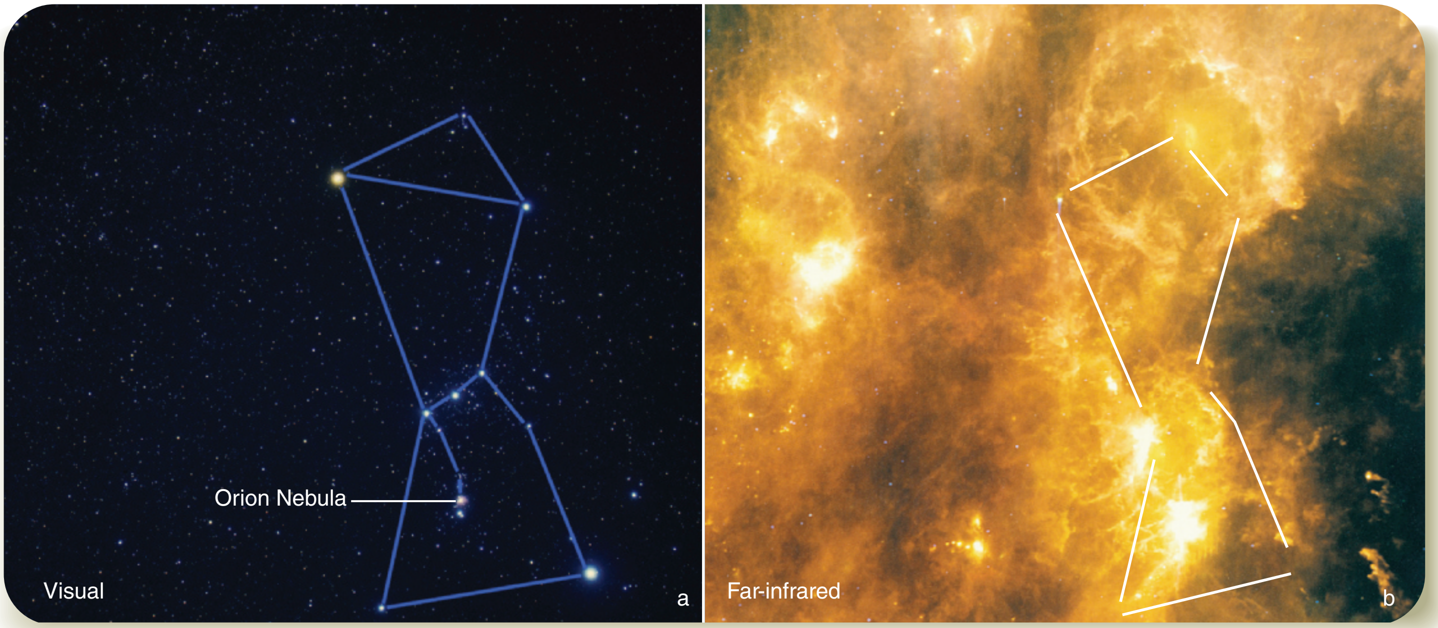
In this chapter, you will see how gravity creates stars from interstellar clouds. You will then learn more about how the flow of energy outward from the insides of stars balances gravity and makes the stars stable, and how nuclear reactions in the cores of stars ultimately supply that energy flow. To understand this story, you will plunge from the cold gas of the interstellar medium into the hot centers of the stars themselves.

14-1 The Interstellar Medium

WHEN PEOPLE USE THE PHRASE "the vacuum of space," they reveal a **Common Misconception** that space is empty. Space is not empty; there is gas and dust between the stars, called the **interstellar medium** (abbreviated, **ISM**). It is not easily detected at wavelengths visible to human eyes, but when observed at other wavelengths it is strikingly complex and beautiful (■ Figure 14-1). The densest clouds of the interstellar medium are where new stars are born.

Observing Interstellar Gas and Dust

One way to investigate properties of the interstellar medium is to measure its effects on the starlight that passes through it. Astronomers observe that distant stars appear fainter than they should for their distances and luminosities (estimated from their spectral types). This effect is called **interstellar extinction.** Distant stars appear not only fainter but also redder than they should for their spectral types. For example, some distant hot O stars actually look red rather than blue. This effect is called **interstellar reddening**

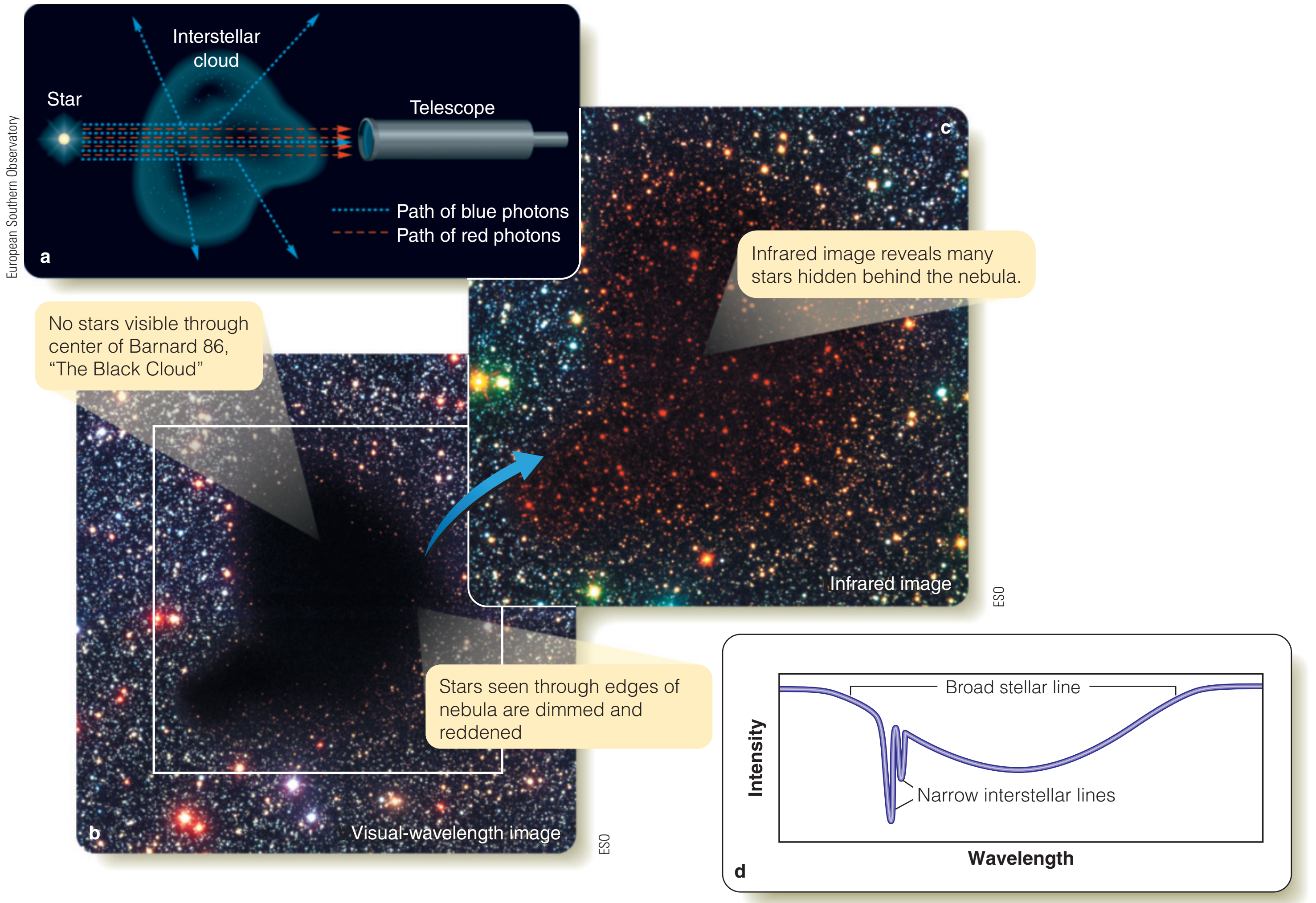


Akira Fujii/David Malin Images

NASA/IPAC courtesy Deborah Levine

■ **Figure 14-1**

(a) At visual wavelengths, the spaces between the stars in and around the constellation Orion seem empty. (b) An infrared image reveals the swirling clouds of the interstellar medium.



■ Figure 14-2

(a) Interstellar reddening makes stars seen through a cloud of gas and dust look redder than they should because shorter wavelength light is more easily scattered. (b) No stars are detected at visual wavelengths through a dense cloud except near the edges. (c) At longer, near-infrared wavelengths, many stars can be detected behind the cloud. (d) Interstellar spectral absorption lines can be recognized because they are very narrow. Here a spectral line produced in the atmosphere of a star is magnified so that it is seen as a broad dip in the plot. In contrast, absorption lines produced by two clouds in the interstellar medium are much narrower.

and is the result of interstellar extinction being stronger for short-wavelength light than for long-wavelength light (■ Figure 14-2). In fact, the interstellar medium is nearly transparent at near-infrared wavelengths, allowing astronomers to look into and through interstellar clouds that are so dense they completely block our view at visual wavelengths (Figure 14-2c).

Another way astronomers detect material in the interstellar medium is by observing **interstellar absorption lines** in the spectra of stars. You can be sure that the interstellar absorption lines are not spectral lines of the stars themselves for several reasons. For example, some distant O stars show calcium and sodium lines, but those stars are too hot to form those lines in their own atmospheres (look back to Chapter 13). Those absorption lines must be produced instead by atoms located between the stars and Earth.

The widths of the interstellar lines also give away their origins. In the atmosphere of a star, the gas is so dense that atoms collide with each other often, disturbing the electron energy levels, and so hot that atoms move rapidly, producing Doppler shifts. Both of those effects broaden (blur) spectral lines from stellar atmospheres. In contrast, much of the interstellar gas is cold, and its density is very low, typically about 1 atom per cubic centimeter. Although the density can be much higher in dense clouds, even there the gas is 10^{14} times less dense than the air you breathe. Because atoms in the interstellar medium move slowly and collide only rarely, interstellar absorption lines are extremely narrow (■ Figure 14-3).

You learned in Chapter 13 that detailed analysis of stellar spectra reveals the composition of stars. In the same way, spectroscopy also tells you the abundances of elements in the interstellar medium. About 70 percent of the mass of interstellar

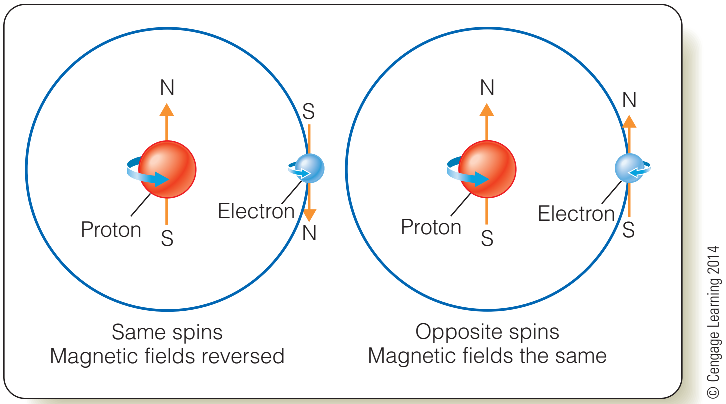


Figure 14-3

Protons and electrons spin and consequently have small magnetic fields. Because they have opposite electrical charges, they have opposite magnetic fields if they spin in the same direction. If they spin in opposite directions, their magnetic fields are parallel. As explained in the text, neutral hydrogen atoms "flipping" from the parallel to the antiparallel state emit radio photons with a wavelength of 21 cm.

gas is hydrogen, 28 percent is helium, and 2 percent is made up of elements heavier than helium. (Look back to Table 7-1 on page 123 and notice that this is the same composition as the sun and other stars.) The atoms are joined together to form molecules in the densest, coldest interstellar clouds, which are called **molecular clouds.** Astronomers have detected over 150 different molecules in the interstellar medium.

Almost all of the material in the interstellar medium is gas, but about 1 percent of the mass is made up of dust particles called **interstellar dust.** Studies of interstellar spectra reveal that the dust particles are composed of carbon, silicates, iron, ice, and organic compounds. (Note that organic compounds are molecules based on multiple linked carbon atoms, but they do not have to originate with living things.) Details of interstellar extinction and reddening show that the particles are almost all less than 1 micron (only 10^{-6} meters, or 1000 nm) in diameter.

You can expect that if interstellar gas atoms are viewed against a dark or cold background rather than in front of the hot bright background of a star's photosphere, those atoms can produce an emission spectrum (look back to Kirchhoff's laws, Chapter 6, page 110). Some parts of the interstellar medium do exactly that. **Interstellar emission lines,** like absorption lines, allow astronomers to study that gas and dust between the stars.

Cold, neutral hydrogen floating in space produces a prominent emission line at a wavelength that can be detected with radio telescopes. This is due to the fact that electrons and protons have tiny magnetic fields because they are charged particles that are also spinning. A hydrogen atom with its electron's magnetic field pointing in the same direction as the proton's magnetic field has slightly more energy than if the electron and proton have magnetic fields aligned in opposite directions. If the electron is spinning so its magnetic field is parallel to the proton's field, it can flip over and spin the other way, releasing the excess energy as a photon with a wavelength of 21 cm. Astronomers can use the resulting hydrogen **21-cm radio line** to map the location of the cold, low-density gas in our galaxy (■ Figure 14-4a).

Infrared emission from the interstellar medium, like radio emission, normally tells you about colder, less excited gas and dust in clouds quietly drifting through space (Figure 14-4b) (look back to Wien's law, Chapter 6, pages 107–108). On the other hand, at optical wavelengths, the interstellar medium is mostly a nuisance, dimming and changing the colors of distant objects viewed through it (Figure 14-4c). Emission from the interstellar medium is also detected at ultraviolet and X-ray wavelengths. In general, high-energy photons are produced by high-energy events, so X-ray observations tell you about very hot, excited gas such as matter expelled by exploding stars (Figure 14-4d).

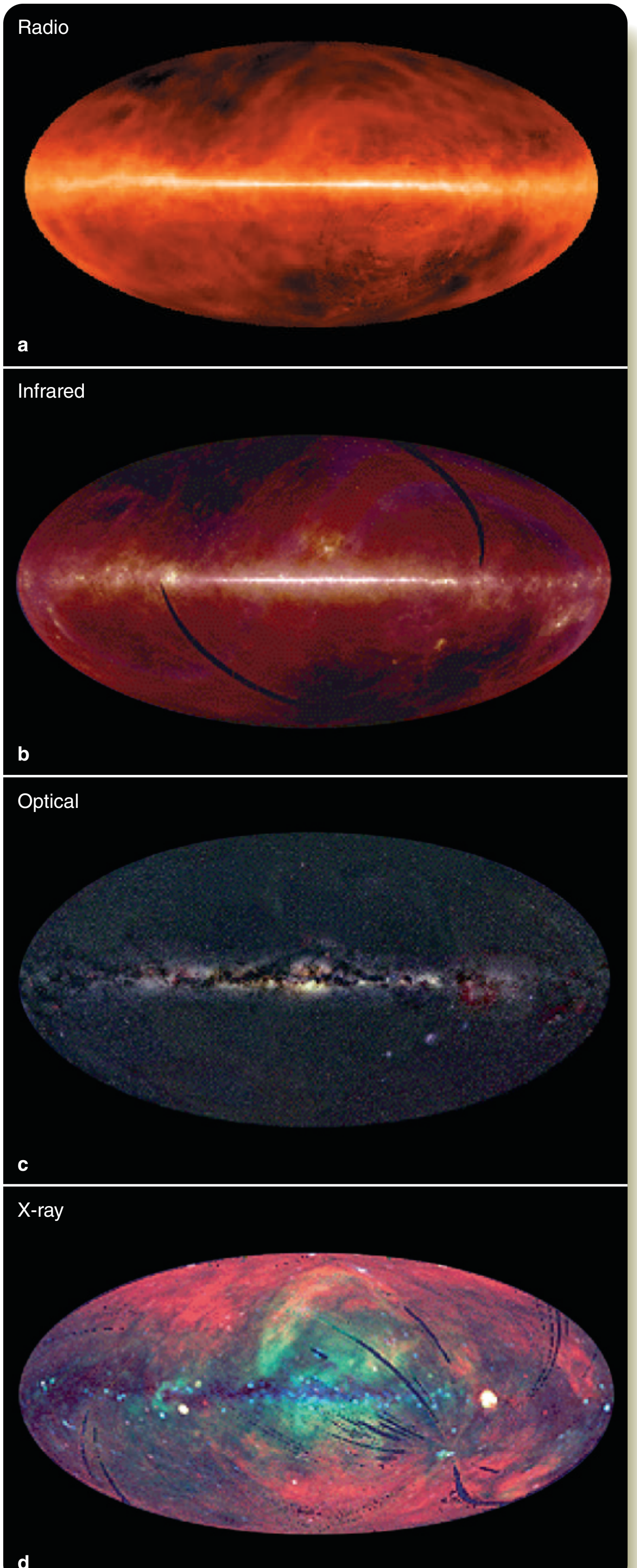
Nebulae

On a cold, clear winter night, Orion hangs high in the southern sky, a large constellation composed of brilliant stars. If you look carefully at Orion's sword, you will see that one of the stars is a hazy cloud (Figure 2-4). A small telescope reveals even more such clouds of gas and dust scattered around the sky. Astronomers refer to these clouds as **nebulae** (singular, **nebula**), from the Latin word for mist or cloud. They are places where the interstellar medium is densest and most easily seen.

Read **Three Kinds of Nebulae** on pages 294–295 and notice three important points and four new terms:

1. *HII regions* (pronounced, *H-two*) are clouds of gas ionized by the ultraviolet radiation from nearby hot stars. They are often called *emission nebulae* because the ionized hydrogen produces visible-wavelength photons that make the regions glow with a characteristic pink color.
2. If nearby stars are not hot enough to ionize the hydrogen in a nebula, it may still be visible as a *reflection nebula,* produced when light is scattered by tiny dust particles mixed in with the gas.
3. At visual wavelengths, *dark nebulae* are visible where dense clouds of gas and dust are silhouetted against background regions filled with stars or bright nebulae.

The different kinds of nebulae that decorate the sky are really similar objects seen in different ways, much like rain clouds in Earth's atmosphere. On a sunny day, a distant cloud looks white, but overhead it blocks sunlight and looks dark. If you could see the same cloud with infrared eyes, you would see it glowing brightly with blackbody radiation. Different kinds of nebulae are regions of the interstellar medium interacting in different ways with the starlight shining on them and through them.



SCIENTIFIC ARGUMENT

What evidence can you cite that there is an interstellar medium? Everything in science is based on evidence, so your argument should discuss observations. First, there is the simple fact that certain parts of the interstellar medium are visible—the nebulae. Detailed analysis of absorption and emission lines show that all of interstellar space is filled by cold, thin gas. Extinction and reddening reveal that some of the material is in the form of tiny dust specks.

Perhaps this is enough evidence to convince someone that the spaces between the stars are not empty, but there is more. Expand your argument. **What do infrared observations reveal about the interstellar gas?**

14-2 Making Stars from the Interstellar Medium

The key to understanding star formation is realizing the relation between young stars and the densest parts of the interstellar medium (■ Figure 14-5). Where you find stars that evidently must have formed recently, there you also find large, dense clouds of gas and dust. For example, some of those interstellar clouds glow brightly because they are illuminated by the hottest, most massive and luminous stars, of spectral types O and B. As you will learn in detail later in this chapter, those kinds of stars pour out such floods of energy that they can last only a few million years, a very short time in astronomical terms, so you must be seeing them pretty near where they were born. O and B stars are therefore one type of signpost pointing to star forming regions. But how can those cold interstellar clouds contract, heat up, and become stars?

Star Birth in Molecular Clouds

As you saw in the previous section, spectra of the interstellar medium show that its chemical composition is basically the same as that of stars. By observing at far-infrared and radio wavelengths, to which interstellar dust is nearly transparent, astronomers can peer inside even dense molecular clouds and find evidence that stars are forming there. In fact, the largest molecular clouds are each massive enough to make thousands of new stars.

■ **Figure 14-4**

These images of our galaxy include the entire sky spread into an oval. (a) 21-cm radio observations reveal many clouds of neutral hydrogen orbiting the center of the galaxy. (b) The infrared image shows the location of clouds of cold interstellar gas and dust made visible by blackbody radiation from dust grains. (c) In the visual-wavelength image, dense, dusty clouds are silhouetted against distant stars. (d) The X-ray image shows the location of very hot gas produced in most cases by the supernova explosions of dying massive stars.

Figure 14-5

A giant molecular cloud collapsed and began forming stars a few million years ago, including three star clusters visible in this image. Intense radiation and outflowing gas from the hottest, most luminous of these stars is eating into the dark molecular cloud and exciting some of its gases to glow. The nebula is located 210,000 ly from Earth in a nearby galaxy.

But such a cloud can form into stars only if small portions of the cloud can be compressed somehow to high density and high temperature. At least four factors resist the compression of an interstellar gas cloud, and those factors must be overcome by gravity before star formation can begin.

First, thermal energy in the gas appears as motion of the atoms and molecules. This thermal motion would make the cloud come apart if gravity were too weak to hold it together. *Second,* the interstellar medium is filled with magnetic fields. Ionized atoms, having an electric charge, cannot move freely through a magnetic field. Although the gas in a molecular cloud is mostly neutral, some ions are included, and that means a magnetic field can exert a force on the gas as a whole. Gravity must overcome the interstellar magnetic field to make the gas contract. *Third,* everything in the universe rotates. As a gas cloud begins to contract, it spins more and more rapidly, just as ice-skaters spin faster as they pull in their arms. This rotation can become so rapid that it resists further contraction of the cloud. *Fourth,* turbulence in the interstellar medium can prevent a cloud from contracting. Many nebulae are observed to be distorted and twisted by strong turbulent currents that can prevent a cloud from contracting (Figure 14-1).

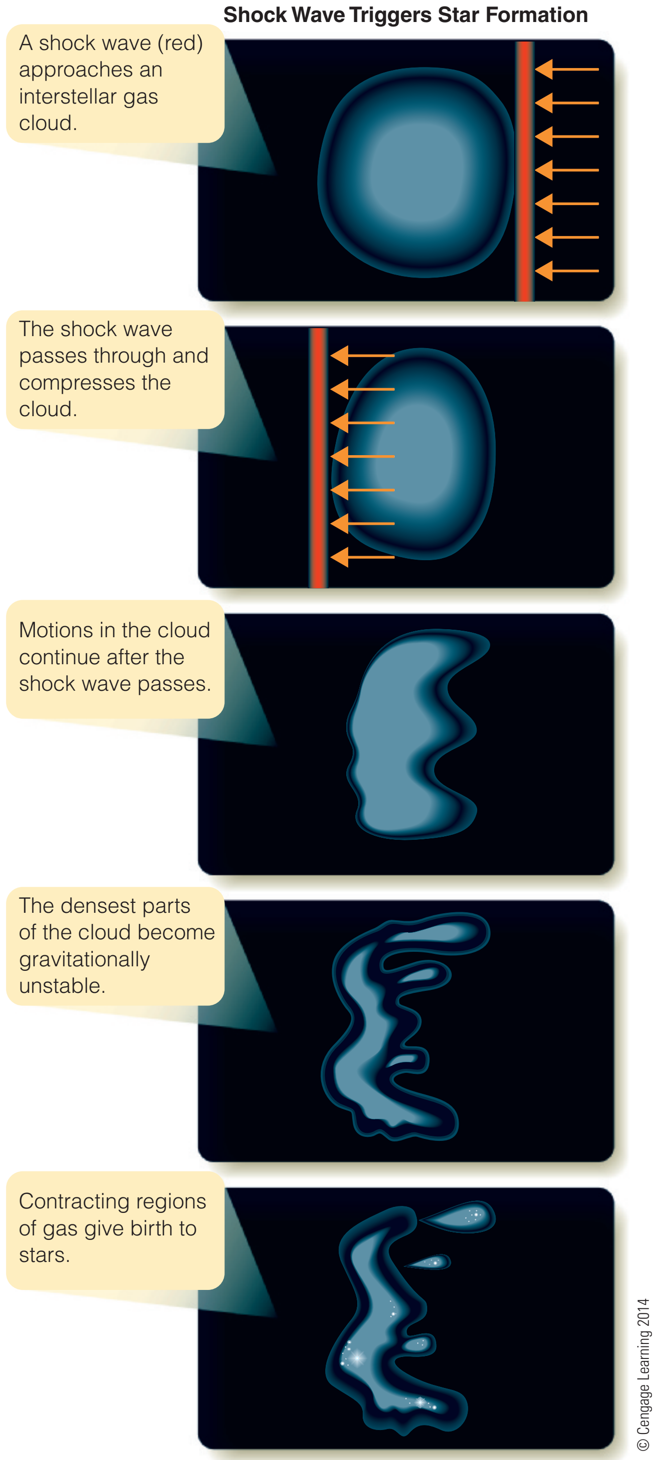


Figure 14-6

In this summary of a computer model, an interstellar gas cloud is triggered into collapsing and forming clusters of new stars by a passing shock wave. The events summarized here might span about 6 million years.

Given these four resisting factors, it may seem surprising that any molecular clouds can contract at all, but both theory and observation suggest that such clouds can be triggered to form stars by effects such as a passing shock wave, sometimes just called a **shock,** the astronomical equivalent of a sonic boom (Figure 14-6).

Three Kinds of Nebulae

1 **Emission nebulae** are produced when a hot star excites the gas near it to produce an emission spectrum. The star must be hotter than about B1 (25,000 K). Cooler stars do not emit enough ultraviolet radiation to ionize the gas. Emission nebulae have a distinctive pink color produced by the blending of the red, blue, and violet Balmer lines. Emission nebulae are also called **HII regions,** following the custom of naming gas with a roman numeral to show its state of ionization. HI is neutral hydrogen, and HII is ionized.

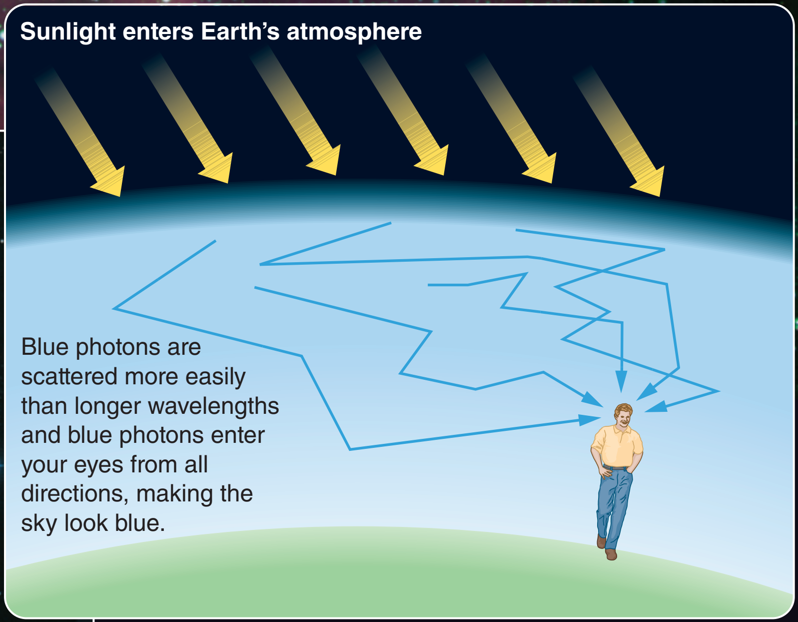
In an HII region, the ionized nuclei and free electrons are mixed. When a nucleus captures an electron, the electron falls down through the atomic energy levels, emitting photons at specific wavelengths. Spectra indicate that the nebulae have compositions much like that of the sun—mostly hydrogen. Emission nebulae have densities of 100 to 1000 atoms per cubic centimeter, better than the best vacuums produced in laboratories on Earth.

Visual-wavelength image

European Southern Observatory

2 A **reflection nebula** is produced when starlight scatters from a dusty nebula. Consequently, the spectrum of a reflection nebula is just the reflected absorption spectrum of starlight. Gas is surely present in a reflection nebula, but it is not excited to emit photons. See image below.

2a Reflection nebulae look blue for the same reason the sky looks blue. Short wavelengths scatter more easily than long wavelengths. See image below.



Reflection nebulae NGC 1973, 1975, and 1977 lie just north of the Orion nebula. The pink tints are produced by ionized gases deep in the nebulae.

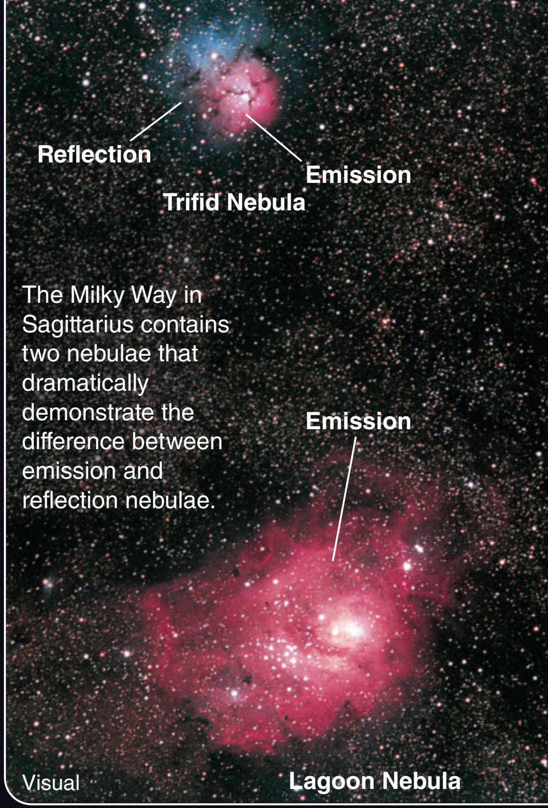
Visual-wavelength image

Australian Astronomical Observatory/David Malin

2b The blue color of reflection nebulae at left shows that the dust particles must be very small in order to preferentially scatter the blue photons. Interstellar dust grains must have diameters ranging from 0.01 mm down to 100 nm or so.

Line art on this page © Cengage Learning 2014

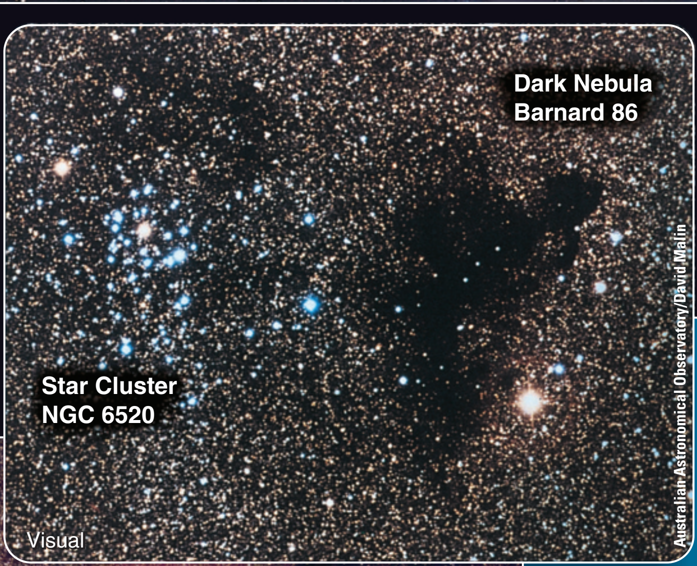
2c In the star cluster called the Pleiades, the hottest stars are B6, not hot enough to ionize hydrogen in the interstellar medium. Instead, the brightest stars produce a reflection nebula as their light is scattered from interstellar dust.



The Milky Way in Sagittarius contains two nebulae that dramatically demonstrate the difference between emission and reflection nebulae.

Daniel Good

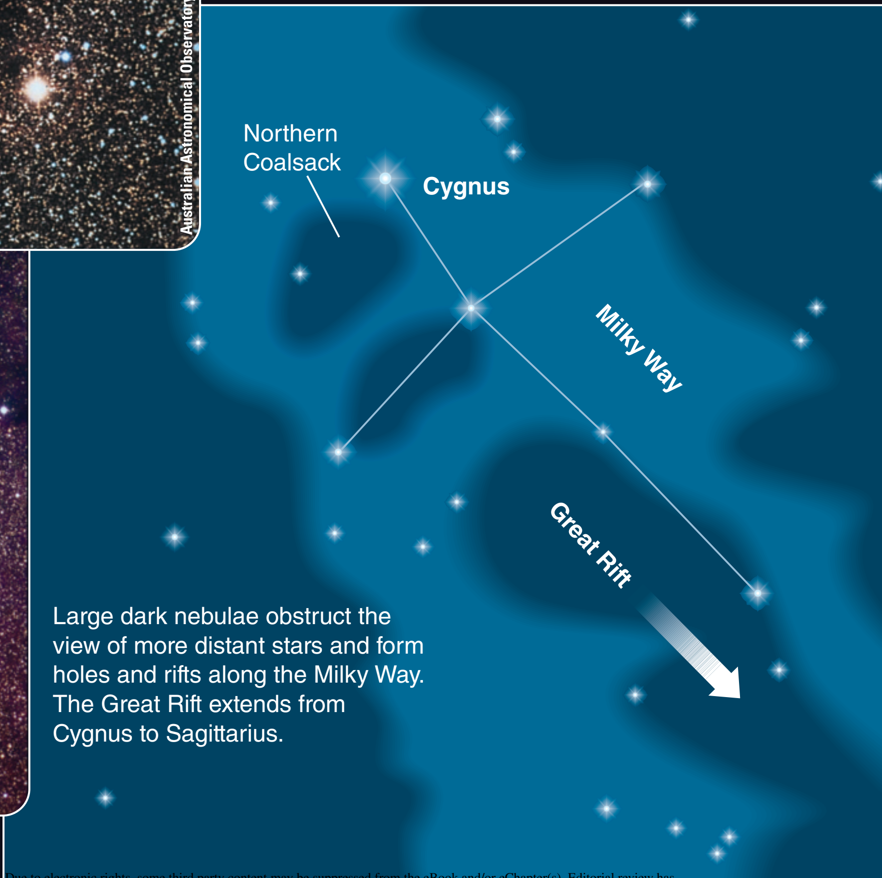
Australian Astronomical Observatory/
David Malin



3 **Dark nebulae** are dense clouds of gas and dust that obstruct the view of more distant stars. Some are generally round, but others are twisted and distorted, as shown at the left, suggesting that even when there are no nearby stars to ionize the gas or produce a reflection nebula, there are breezes and currents pushing through the interstellar medium.

Australian Astronomical Observatory/
David Malin

Twisted by intense light from nearby stars, this dark nebula is visible because it obscures more distant stars.



Large dark nebulae obstruct the view of more distant stars and form holes and rifts along the Milky Way. The Great Rift extends from Cygnus to Sagittarius.

Line art on this page © Cengage Learning 2014

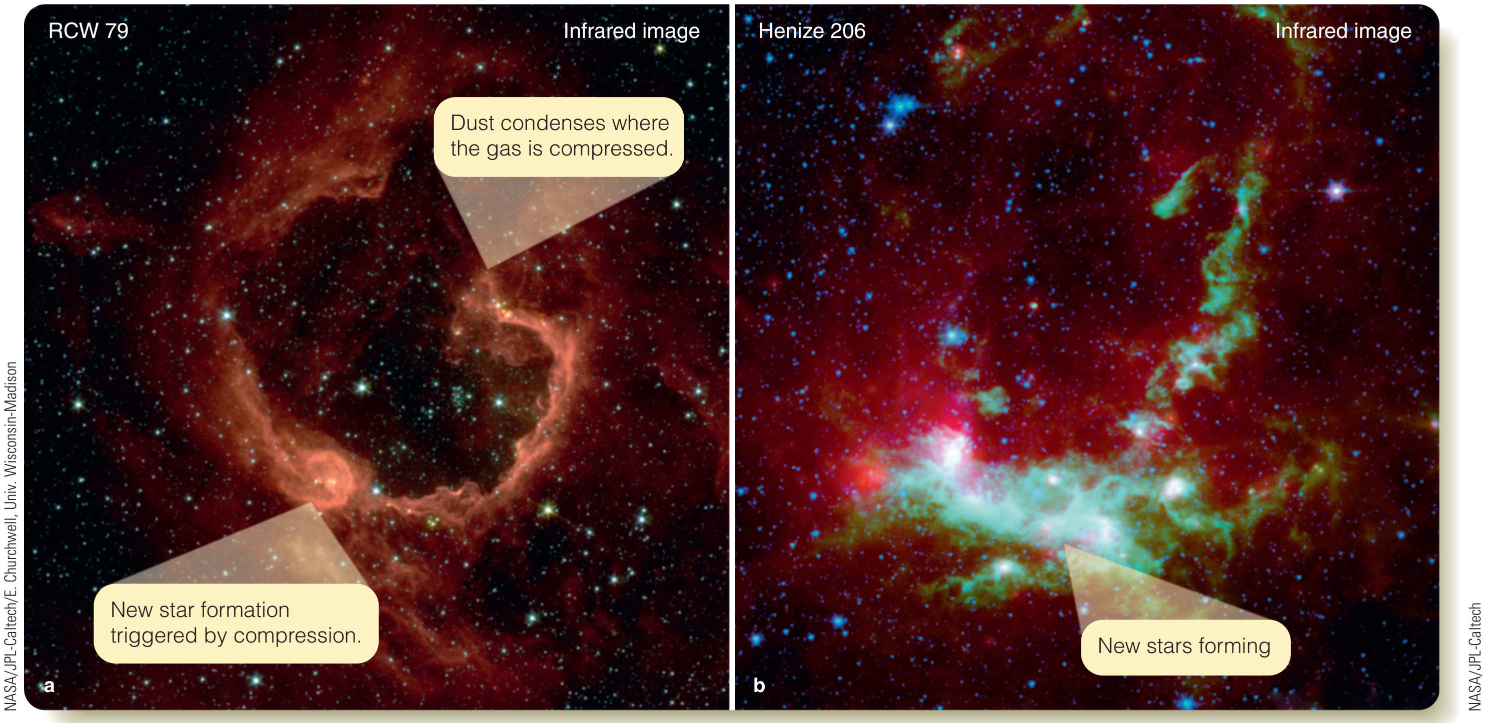


Figure 14-7

(a) Nebula RCW 79 is a bubble of intensely hot gas about 70 ly in diameter. It was produced over the last million years as hot gas and radiation flowed away from massive, hot stars near the center of the bubble. (b) Millions of years ago, a massive star exploded as a supernova in the upper left quarter of this image, and hot gas from that explosion is now pushing into cooler gas and causing the condensation of dust, which radiates strongly at infrared wavelengths.

During such a triggering event, some regions of the large cloud can be compressed to such high densities that the resisting factors can no longer oppose gravity, and star formation begins.

Shock waves that can trigger star formation are common in the interstellar medium. For example, gas flows of several sorts produce shocks when they push into colder, denser interstellar matter. The ignition of very hot stars ionizes nearby gas and causes it to flow rapidly away (■ Figure 14-7a), and new stars of all types seem to emit strong winds and jets while they are forming. Supernova explosions (see Chapter 15) create powerful shock waves that rush through the interstellar medium (Figure 14-7b). Each of those kinds of flows can produce shocks. Collisions between molecular clouds can compress parts of the clouds and also cause star formation. Evidently another type of star formation trigger is the spiral pattern of our Milky Way Galaxy (look back to Figure 1-11). Theory and observations both indicate that spiral arms are shock waves that travel around the disk of the galaxy like the moving hands of a clock (see Chapter 17). As a cloud passes through a spiral arm, the cloud can be compressed, and star formation may begin. Astronomers have found regions of star formation in which each of these triggering processes can be identified.

A single **giant molecular cloud** containing a million solar masses does not contract to form a single humongous star. Evidently, collapsing clouds break into fragments, and the densest parts form a number of dense cores. Exactly why clouds divide into fragments isn't fully understood, but their rotation, magnetic fields, and turbulence apparently play important roles. Whatever the reason, when a giant cloud of interstellar matter contracts, it forms many newborn stars simultaneously.

Protostars

You can understand how low-density clouds of interstellar gas might contract to become as dense as stars, but how does the cold gas become hot enough to be a star? The answer, once again, is gravity.

To see how gravity can heat the gas, shift your attention to one dense cloud core destined to become a single star. Once that small cloud of gas begins to contract, gravity causes the atoms to fall toward the center, gaining speed as they go. Astronomers refer to this early stage in the formation of a star as **free-fall collapse.** By the time the atoms have fallen most of the way to

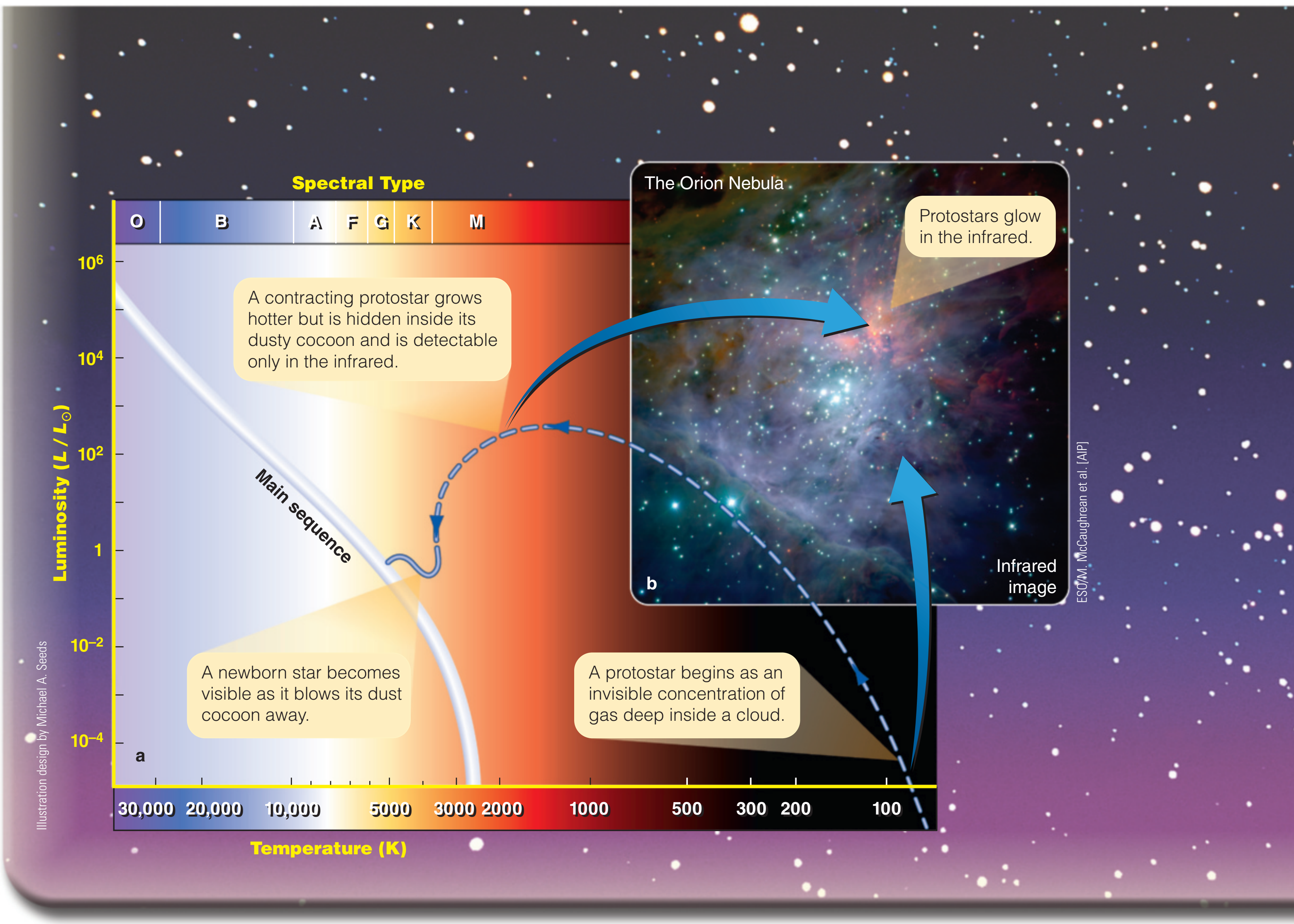
the center of the cloud, they are traveling rapidly. The atoms begin to collide with one another and convert their inward motion to jumbled, random motion of thermal energy, causing the temperature of the gas to increase.

When the contracting core becomes hot enough, it produces short-wavelength radiation to which the core is not transparent. Then the core's thermal energy can no longer escape, and it enters a new stage of slow, rather than free-fall, contraction. Although the term is used rather loosely by astronomers, a **protostar** can be defined as a stellar-mass object that is compressed and hot enough to be opaque to the radiation it is emitting internally but not yet compressed and hot enough to be able to generate energy by nuclear fusion, as would a main-sequence star. Instead, a protostar shines with energy released gravitationally, from a combination of matter falling inward to the protostar plus slow contraction of the protostar itself.

The protostar continues growing deep inside a surrounding cloud of cold, dusty gas. Those enveloping clouds have been called **cocoons** because they hide the forming protostar from view. That is especially true for observations at short wavelengths. The cocoon absorbs the protostar's radiation and, growing warm, reradiates the energy as longer-wavelength infrared radiation than the radiation produced by the protostar. In other words, if you observe at short infrared wavelengths, you may be able to observe the protostar itself; at longer wavelengths, you might be able to observe its cocoon. In either case, what you detect would be cooler than a star but much larger than the sun and would therefore be a very luminous infrared source, with properties placing it in the upper right part of the H–R diagram (■ Figure 14-8).

■ Figure 14-8

(a) This H–R diagram has been extended to very low temperatures to show schematically the contraction of a dim, cool protostar. (b) Protostars are normally invisible at visual wavelengths because they are cool objects and are deep inside dusty clouds of gas, but they are detectable at infrared wavelengths.



How Do We Know? 14-1

Separating Facts from Hypotheses

When scientists disagree, what do they debate? The fundamental work of science is testing theories by comparing them with facts. The facts are evidence of how nature works, and they represent reality. Hypotheses are attempts to explain how nature works. Scientists are very careful to distinguish between the two.

Scientific facts are those observations or experimental results of which scientists are confident. Ornithologists might note that fewer mountain thrushes are returning each spring to a certain mountain valley. Counting bird populations reliably is difficult and requires special techniques, but if the scientists made the observations correctly, they can be confident of their result and treat it as a fact.

To explain the declining population of thrushes, the scientists might consider a number of hypotheses, such as global warming or chemical pollution in the food chain. The ornithologists are free to combine or adjust their hypotheses to better explain the bird migration, but they are not free to adjust their facts. Scientific facts are the hard pebbles of reality that can't be changed.

New facts can aggravate debates that are already politically charged. The declining number of mountain thrushes, for example, could be unwelcome news because addressing the root problem might cost taxpayer dollars or hurt local business interests. Nonscientists sometimes debate an issue by trying to adjust or even deny the facts, but scientists are not free to ignore a fact because it is unpopular or inconvenient. Scientists debate an issue by arguing about which hypotheses best applies or how a hypothesis could be adjusted to fit the observed facts, but, once established, the facts themselves are not in question.

Whether scientists are measuring the density of an emission nebula or the size of a bird population, the final data become the reality against which hypotheses are tested. When Galileo said we should "read the book of nature," he meant we should consult reality as the final check on our understanding.

Michael A. Seeds

In science, evidence is made up of facts, which could range from precise numerical measurements to the observation of the shape of a flower.

The theory of star formation takes you into an unearthly realm filled with unfamiliar processes and objects. How can anyone really know how stars are born? Although the theory of star formation, like all scientific theories, can never be absolutely proven, scientists have strong confidence in it because of repeated testing by observation (**How Do We Know? 14-1**). There is plenty of evidence that star formation is a continuous process; you can be sure that stars are being born right now.

The Orion Nebula: Evidence of Star Formation

In astronomy, evidence means observations. Consequently, you can ask what observations confirm the hypotheses about star formation. Protostars are not easy to observe. Astronomers calculate that the protostar stage is less than 1 percent of the lifetime of a star like the sun; and, although that is a long time in human terms, you cannot expect to catch many stars in the protostar stage. Furthermore, protostars form deep inside clouds of dusty gas that block your view at visible wavelengths, and, being cooler than full-fledged stars, they produce radiation primarily at longer wavelengths. Astronomers must therefore depend on observations at infrared and radio wavelengths to search for protostars in their natural environment.

At the center of the Orion Nebula, also known as Messier 42, lie four brilliant blue-white O and B stars known as the Trapezium, the brightest in a cluster of a few hundred stars. Surrounding those stars are the glowing filaments of a nebula more than 8 pc across. Like a great thundercloud illuminated from within, the churning currents of gas and dust suggest immense power.

The significance of the Orion Nebula lies hidden, figuratively and literally, beyond the visible nebula. This region is ripe with star formation. Read **Star Formation in the Orion Nebula** on pages 300–301 and note four important points and one new term:

1. The nebula you see is only a small part of a vast, dusty molecular cloud. You see the nebula because the larger stars born within it have ionized the gas and driven it outward, breaking out of the molecular cloud. Visual and infrared observations reveal small dusty clouds of gas called *Bok globules* that may be in the process of forming stars.

2. A single very hot and short-lived O star, one of the stars in the Trapezium cluster, is almost entirely responsible for

producing the ultraviolet photons that ionize the gas and make the nebula glow. This star is so massive that it has already become a main sequence star while its smaller siblings are still protostars.

3 Infrared observations reveal clear evidence of active star formation deeper in the molecular cloud behind the visible nebula.

4 Many of the young stars in the Orion Nebula are surrounded by disks of gas and dust.

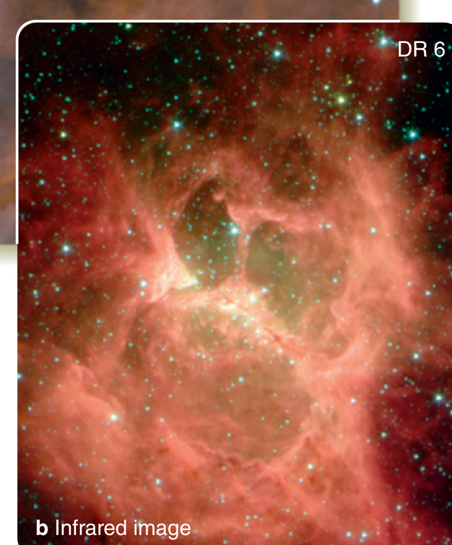
You should not be surprised to find star formation in Orion. The constellation is a brilliant pattern in the winter sky, marked by bright blue stars. Although those stars are hundreds of parsecs away, they appear bright because they are tremendously luminous and cannot live more than a few million years. Their astronomically short lifetimes means they must have been born somewhere near where you see them now. Many other nebulae have properties like the Orion Nebula, showing that they are also sites of star formation.

Contagious Star Formation

Not only can astronomers locate evidence of star formation, but they also have found evidence that star formation can stimulate more star formation. If a gas cloud produces massive stars, those massive stars ionize the gas nearby and drive it away. Where the intense radiation and hot gas pushes into surrounding gas, it can compress the gas and trigger more star formation. Like a brush fire moving through the interstellar medium, star formation can spread itself especially by the actions of massive stars, and astronomers have located the remains of such episodes (■ Figure 14-9). Of course, lower-mass stars also form in the process, but they cannot trigger further star formation because they are not hot enough to ionize and drive away large amounts of gas, nor do they explode as supernovae.

The history of star formation in the constellation of Orion is written in its stars. Judging from the luminosities and estimated masses of the most prominent examples, the stars at Orion's west shoulder are less than 12 million years old, while the stars of Orion's

NASA/ESA/The Hubble Heritage Team/AURA/STScI



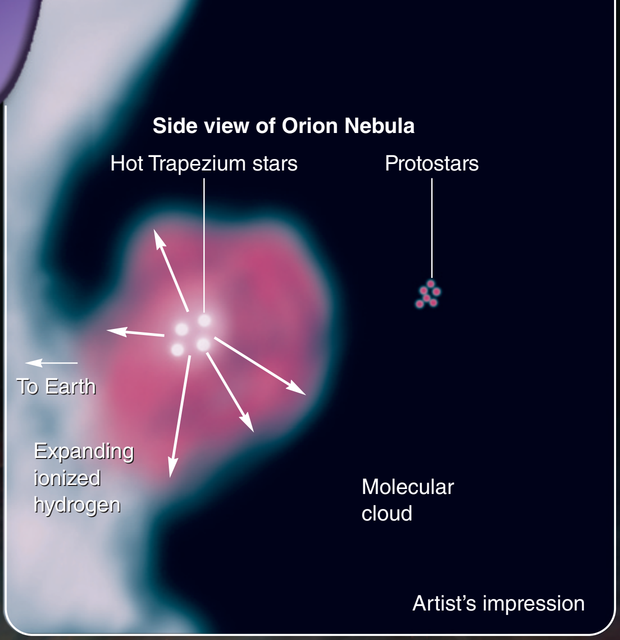
NASA/JPL-Caltech/S. Carey

■ Figure 14-9

(a) The blast of light and ultraviolet radiation from an earlier generation of massive stars has compressed neighboring gas and triggered the formation of more stars. Those stars are now triggering the birth of a third generation at right. (b) Nebula DR 6, dubbed the Galactic Ghoul, is about 3.5 ly in diameter. It was formed by the hot winds and radiation from 10 massive stars at its center that are inflating the nebula, compressing nearby gas, and triggering more star formation.

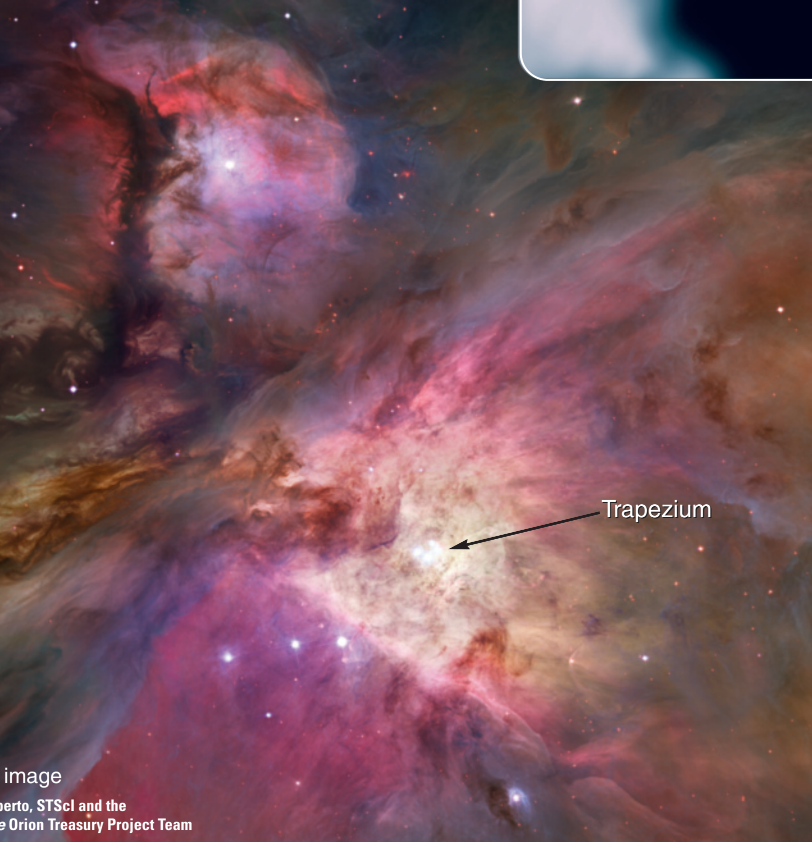
Star Formation in the Orion Nebula

1 The visible Orion Nebula shown below is a pocket of ionized gas on the near side of a vast, dusty molecular cloud that fills much of the southern part of the constellation Orion. The molecular cloud can be mapped by radio telescopes. To scale, the cloud would be many times larger than this page. As the stars of the Trapezium were born in the cloud, their radiation has ionized the gas and pushed it away. Where the expanding nebula pushes into the larger molecular cloud, it is compressing the gas (see diagram at right) and may be triggering the formation of the protostars that can be detected at infrared wavelengths within the molecular cloud.



Hundreds of stars lie within the nebula, but only the four brightest, those in the Trapezium, are easy to see with a small telescope. A fifth star, at the narrow end of the Trapezium, can be visible on nights of good seeing.

The cluster of stars in the nebula is less than 2 million years old. This means the nebula is similarly young.

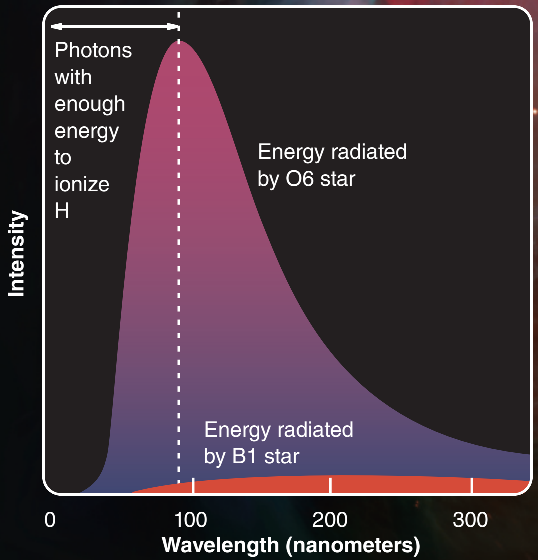


Visual-wavelength image

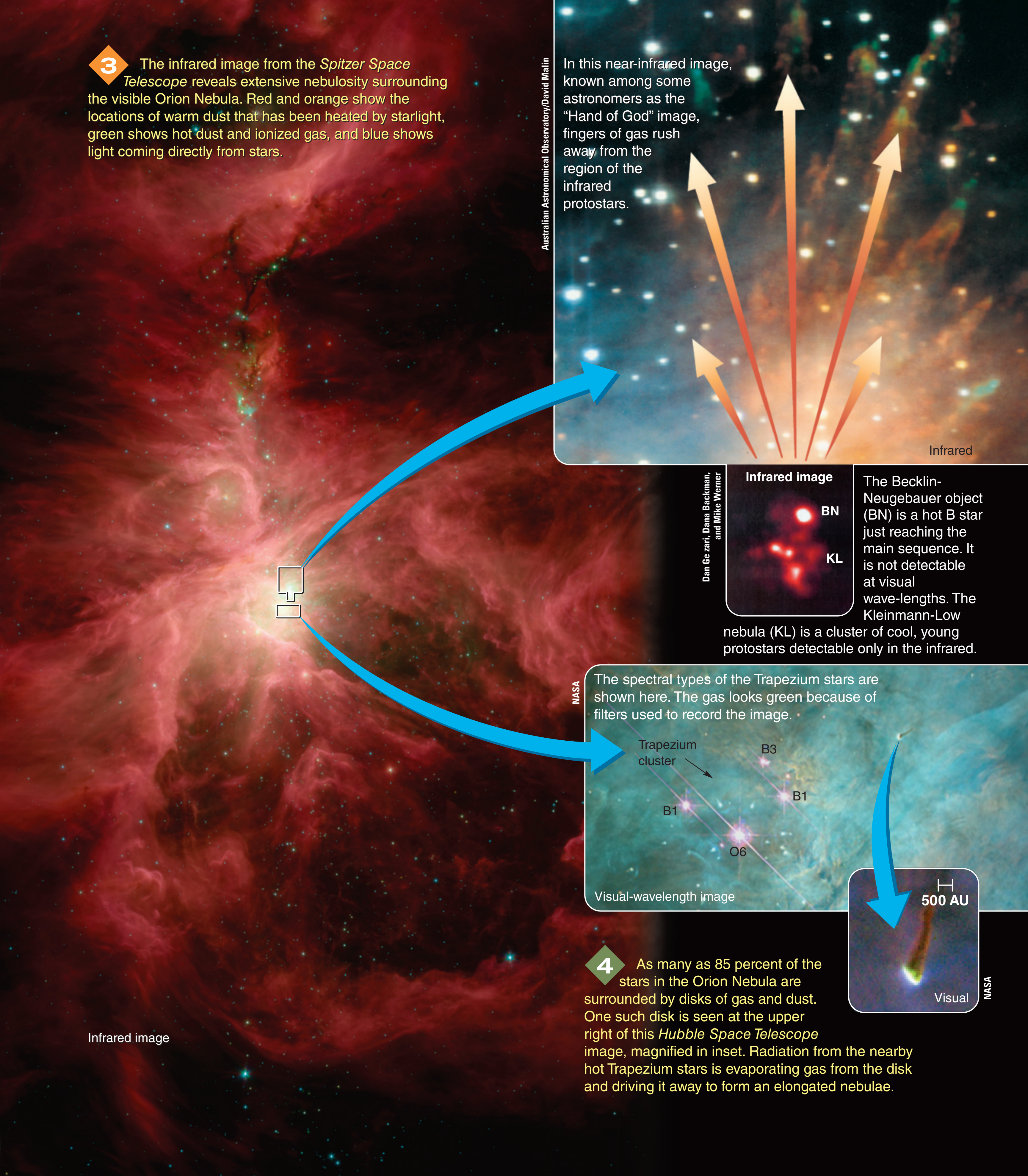
Credit: NASA, ESA, M. Robberto, STScI and the *Hubble Space Telescope* Orion Treasury Project Team

The near-infrared image above reveals more than 50 low-mass, cool protostars.

Small dark clouds called **Bok globules,** named after astronomer Bart Bok, are found in and near star-forming regions. The one pictured above is part of nebula NGC 1999 near the Orion Nebula. Typically about 1 light-year in diameter, they contain from 10 to 1000 solar masses.



2 Of all the stars in the Orion Nebula, only one is hot enough to ionize the gas. Only photons with wavelengths shorter than 91.2 nm can ionize hydrogen. The second-hottest stars in the nebula are B1 stars, and they emit little of this ionizing radiation. The hottest star, however, is an O6 star 30 times the mass of the sun. At a temperature of 40,000 K, it emits plenty of photons with wavelengths short enough to ionize hydrogen. Remove that one star, and the nebula would turn off its emission.


3 The infrared image from the *Spitzer Space Telescope* reveals extensive nebulosity surrounding the visible Orion Nebula. Red and orange show the locations of warm dust that has been heated by starlight, green shows hot dust and ionized gas, and blue shows light coming directly from stars.
Infrared image
Australian Astronomical Observatory/David Malin
In this near-infrared image, known among some astronomers as the "Hand of God" image, fingers of gas rush away from the region of the infrared protostars.
Infrared
Dan Ge zari, Dana Backman, and Mike Werner
Infrared image
BN
KL
The Becklin-Neugebauer object (BN) is a hot B star just reaching the main sequence. It is not detectable at visual wave-lengths. The Kleinmann-Low nebula (KL) is a cluster of cool, young protostars detectable only in the infrared.
NASA
The spectral types of the Trapezium stars are shown here. The gas looks green because of filters used to record the image.
Trapezium cluster
B3
B1
B1
O6
Visual-wavelength image
500 AU
Visual
NASA
4 As many as 85 percent of the stars in the Orion Nebula are surrounded by disks of gas and dust. One such disk is seen at the upper right of this *Hubble Space Telescope* image, magnified in inset. Radiation from the nearby hot Trapezium stars is evaporating gas from the disk and driving it away to form an elongated nebulae.

belt are at most 8 million years old. The stars of the Trapezium at the center of the Great Nebula are no older than 2 million years. Star formation has swept across Orion from northwest to southeast, apparently beginning near Orion's west shoulder. The massive stars that formed there probably triggered the formation of the stars in Orion's belt. That star formation episode may then have caused the formation of the new stars you see today in the Great Nebula.

SCIENTIFIC ARGUMENT

What did Orion look like to the ancient Egyptians, to the first humans, and to the dinosaurs?

Scientific arguments can do more that support a theory; they can change the way you think of the world around you. The Egyptian civilization had its beginning only a few thousand years ago, and that is not very long in terms of the history of Orion. The stars you see in the constellation are luminous and young, but they are a few million years old, so the Egyptians saw the same constellation you see. (They called it Osiris.) Even the Orion Nebula hasn't changed very much in a few thousand years, and Egyptians viewed it in the dark skies along the Nile and perhaps wondered about it.

Our oldest humanoid ancestors lived about 4 million years ago, and that was about the time that the youngest stars in Orion were forming. These earliest ancestors may have looked up and seen some of the stars you see, but some stars have formed since that time. The Great Nebula is excited by the Trapezium stars, and they are no more than about 2 million years old, so our earliest human relatives probably didn't see the Great Nebula.

The dinosaurs saw something quite different. The last of the dinosaurs died about 65 million years ago, long before the birth of the brightest stars in Orion. The dinosaurs, had they the brains to appreciate the view, might have seen bright stars along the Milky Way, but they didn't see Orion. All of the stars in the sky are moving through space, and the sun is orbiting the center of our galaxy. Over many millions of years, the stars move appreciable distances across the sky. The night sky above the dinosaurs contained totally different star patterns.

The Orion Nebula is the product of a giant molecular cloud, but such a cloud can't continue spawning new stars forever. Focus your argument to answer the following: **What processes limit star formation in a molecular cloud?**

14-3 Young Stellar Objects and Protostellar Disks

When most of the material in a protostar's cocoon has fallen inward or been driven away, the protostar will no longer be quite so hidden. The locations in the H–R diagram of protostars that have become detectable at visible wavelengths because their cocoons have disappeared is called the **birth line** (■ Figure 14-10). Once a star crosses the birth line and becomes visible, it continues to contract and move toward the main sequence at a pace that depends on its mass. Stars in this late stage of formation are sometimes called **Young Stellar Objects (YSOs)** or pre-main-sequence stars, to distinguish them from earlier protostellar stages.

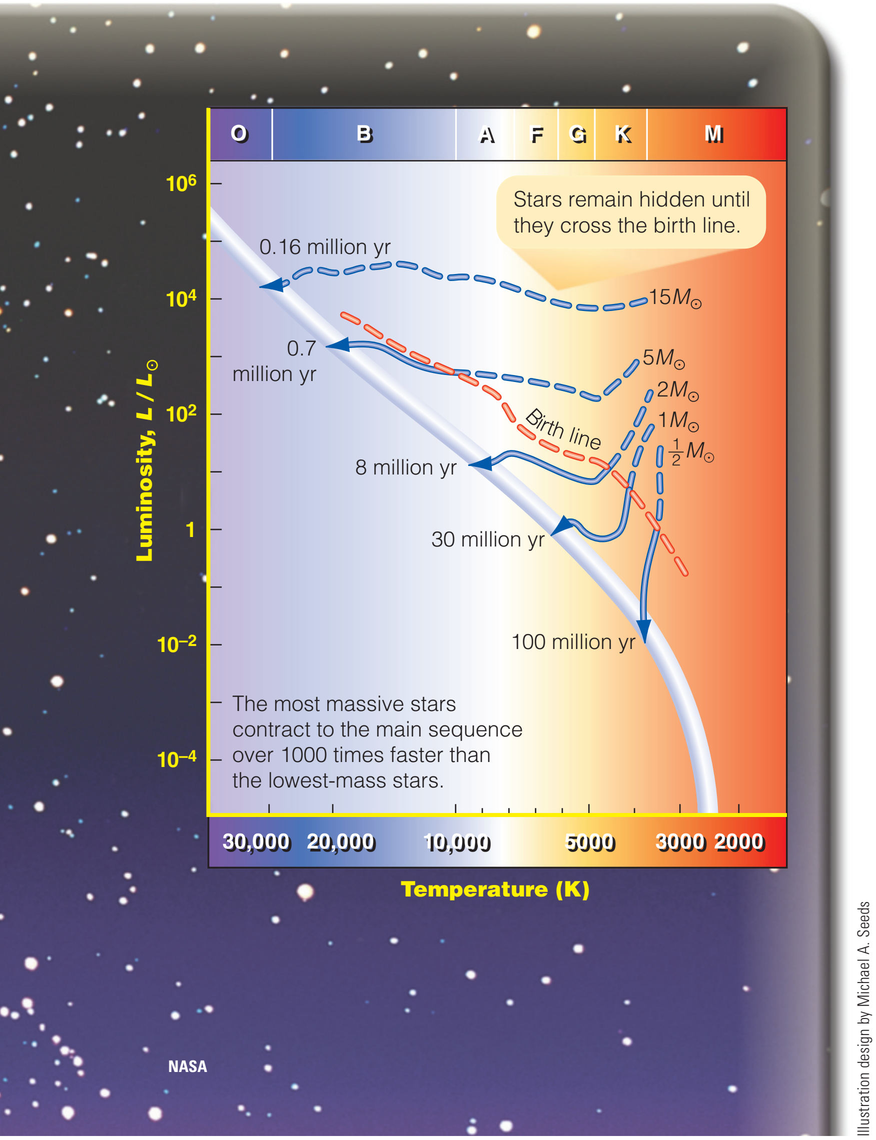
The more massive a star, the stronger its gravity and therefore the more rapid will be its contraction. Astronomers calculate that the sun took about 30 million years from when its cloud began contracting until it became a main-sequence star, but a 30-solar-mass star would take only 30,000 years. A 0.2-solar-mass star needs as much as 1 billion years to finish forming and reach the main sequence.

As a contracting molecular cloud core becomes a protostar, the cloud rotates faster and faster. The rapidly spinning core of the cloud must flatten into a spinning disk like a blob of pizza dough spun into the air. Gas from the envelope that has lost some of its orbital velocity through collisions can sink directly to the center of the cloud, where the protostar grows more massive, surrounded by the disk. Other in-falling material can add to the disk and then move within the disk plane toward the star (■ Figure 14-11).

Theoretically, for a protostar to have no surrounding disk at all, its original cloud would need to have zero rotation, which is very unlikely. Observationally, astronomers find that the majority of protostars appear to be surrounded by disks. The disks that form around protostars are called **protostellar disks,** and they are important because, as you learned in Chapter 8, astronomers conclude that Earth and the other planets of our solar system formed in such a disk around the protosun 4.6 billion years ago. Because most protostars have protostellar disks, it seems likely that most stars have planetary systems.

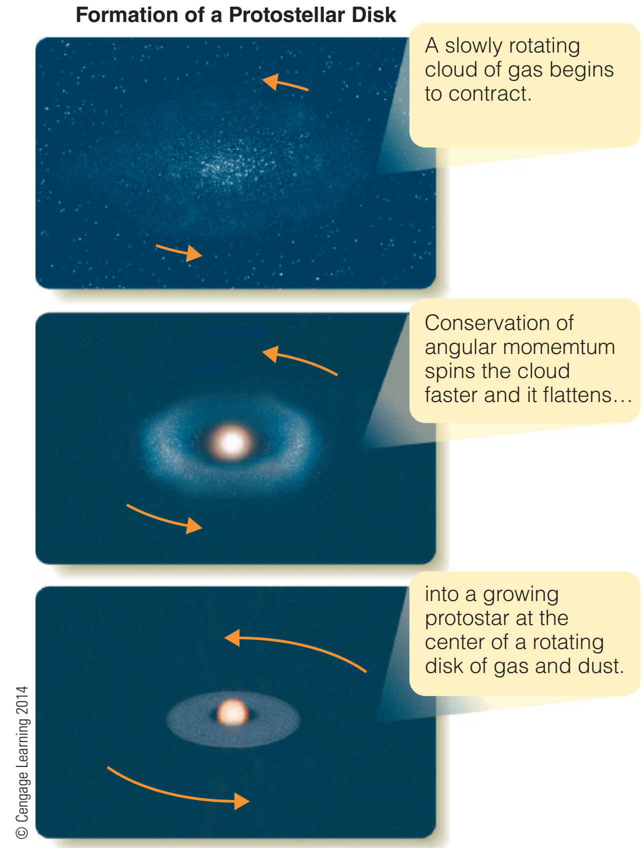
Read **Observations of Young Stellar Objects and Protostellar Disks** on pages 304–305 and notice that it makes four important points and introduces three new terms:

1. Star formation regions contain types of stars so young they must have formed recently. *T Tauri* stars, for example, are a type of star understood to be still in the slowly contracting phase.
2. In the H–R diagram, newborn stars lie between the birth line and the main sequence—just where you would expect to find stars that have recently lost their gas and dust cocoons. Young stars tend to have very active, hot coronae and chromospheres, making them prominent X-ray sources.
3. Observations provide clues about the effects of disks of gas and dust around protostars. These disks evidently cause *bipolar flows* that push into the surrounding interstellar medium and produce glowing blobs called *Herbig-Haro objects.*
4. Finally, notice that images have been obtained of a few protostellar disks. Such disks cannot last very long, and are clear evidence that the stars within them are very young. Astronomers have abundant evidence that these disks are the sites of planet formation.



■ **Figure 14-10**

The more massive a protostar is, the faster it contracts. A 1-$M_{\odot}$ star requires 30 million years to reach the main sequence. (Recall that $M_{\odot}$ means "solar mass.") The dashed line is the birth line, where contracting protostars first become visible as they dissipate their surrounding clouds of gas and dust. Compare with Figure 14-8, which shows the evolution of a protostar of about 1 $M_{\odot}$ as a dashed line up to the birth line and as a solid line from the birth line to the main sequence.



■ **Figure 14-11**

The rotation of a contracting gas cloud forces it to flatten into a disk, and the protostar grows at the center.

In some cases, dark disks of gas and dust are clearly visible around newborn stars (■ Figure 14-12), but it isn't clear how the protostellar disk produces jets. Certainly the contracting, spinning disk contains tremendous energy, and theorists suspect that magnetic fields become twisted tightly around the disk. Exactly how those fields eject hot gas away from the stars is not yet clear, but the presence of a confining disk is thought to explain why material flows out in two streams, along the axis of rotation. The detection of these jets was one of the first pieces of evidence that substantial disks often surround protostars.

Some evidence of stellar youth is more subtle. An **association** is a widely distributed star cluster that is not held together by its own gravity—its stars wander away as the association ages. It is not clear why some gas clouds give birth to compact star clusters held together by their own gravity and others give birth to larger associations not bound together. You can conclude, however, that associations must consist of young stars because the stars wander apart astronomically quickly. The constellation Orion, a known region of star formation, is filled with T Tauri stars in a **T association.** T Tauri stars are relatively low-mass pre-main-sequence objects ranging from roughly 0.5 to 2 solar masses. **OB associations,** fairly obviously, are extended groups of more massive O and B stars. Associations of stars that are journeying away from each other as we watch are clear evidence of recent star formation.

Observations of Young Stellar Objects and Protostellar Disks

1 The nebula around the star S Monocerotis is bright with hot stars. Such stars live short lives of only a few million years, so they must have formed recently. Such regions of young stars are common. The entire constellation of Orion is filled with young stars and clouds of gas and dust.

Nebulae containing young stars usually contain **T Tauri** stars. These stars fluctuate irregularly in brightness, and many are bright in the infrared, indicating that they are surrounded by dust clouds and in some cases by dust disks. Doppler shifts show that gas is flowing away from many T Tauri stars. The T Tauri stars are evidently newborn stars just blowing away their dust coccoons. They have estimated ages ranging from 100 thousand to 100 million years. Spectra of T Tauri stars show signs of an active chromosphere as you might expect from young, rapidly rotating stars with powerful dynamos and strong magnetic fields.

Visual-wavelength image

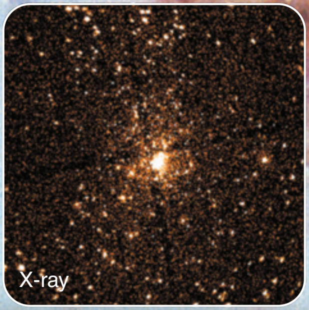
Visual

Infrared

Davide De Martin - http://www.skyfactory.org

NASA/JPL-Caltech/W. Reach

2 The Elephant Trunk (above) is a dark nebula compressed and twisted by radiation and winds from a luminous star beyond the left edge of this image. Infrared observations reveal that it contains six protostars (pink objects, lower edge) not detected at visual wavelengths.

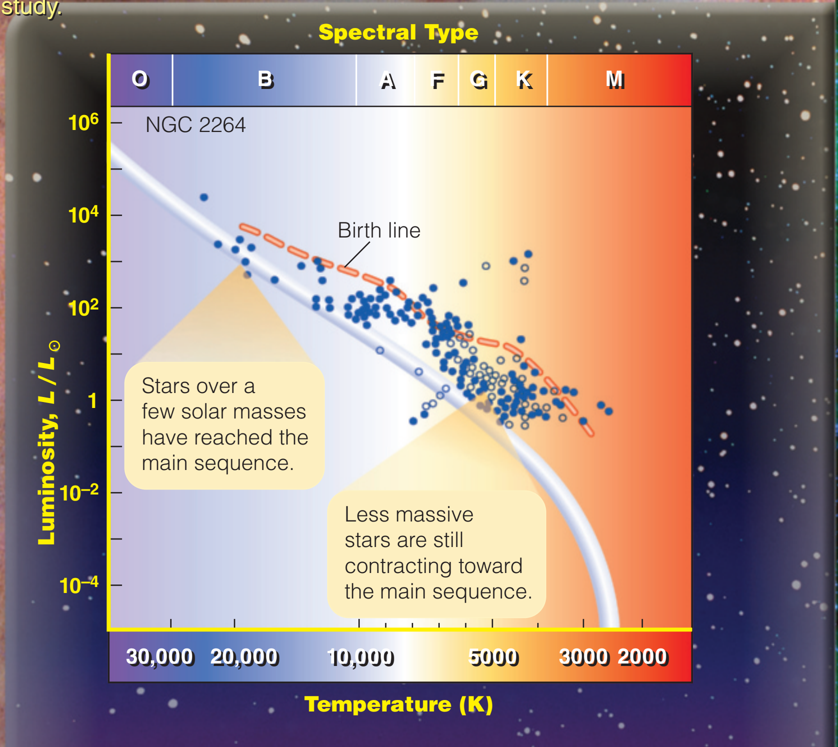


X-ray

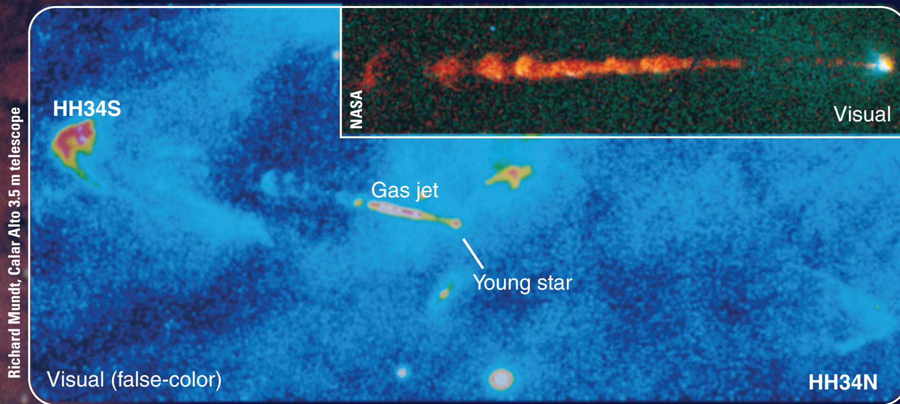
NASA/PSU

Roughly 1000 young stars with hot chromospheres appear in this X-ray image of the Orion Nebula.

3 The star cluster NGC 2264, embedded in the nebula on this page, is only a few million years old. Lower-mass stars have not yet reached the main sequence, and the cluster contains many T Tauri stars (open circles), which have properties above and to the right of the main sequence, near the birth line. The faintest stars in the cluster were too faint to be observed in this study.

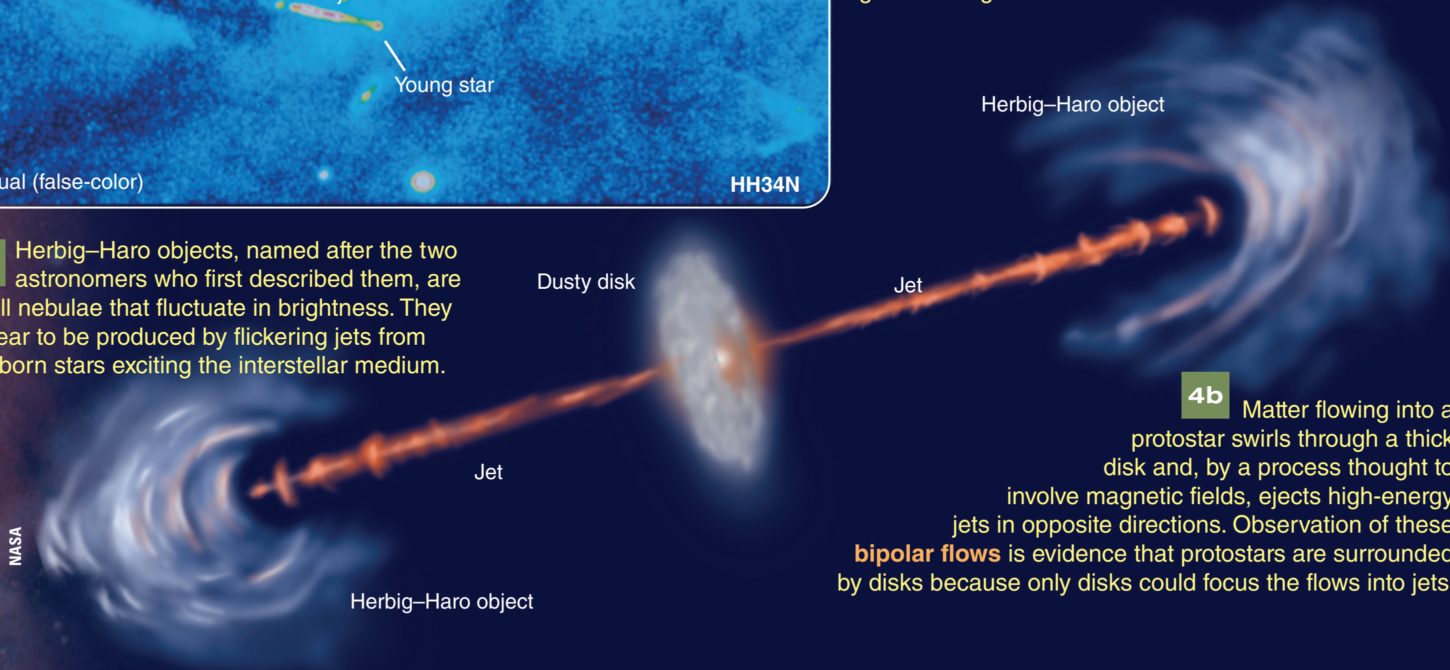


Line art on this page © Cengage Learning 2014; images: NASA/*Hubble* Heritage Project

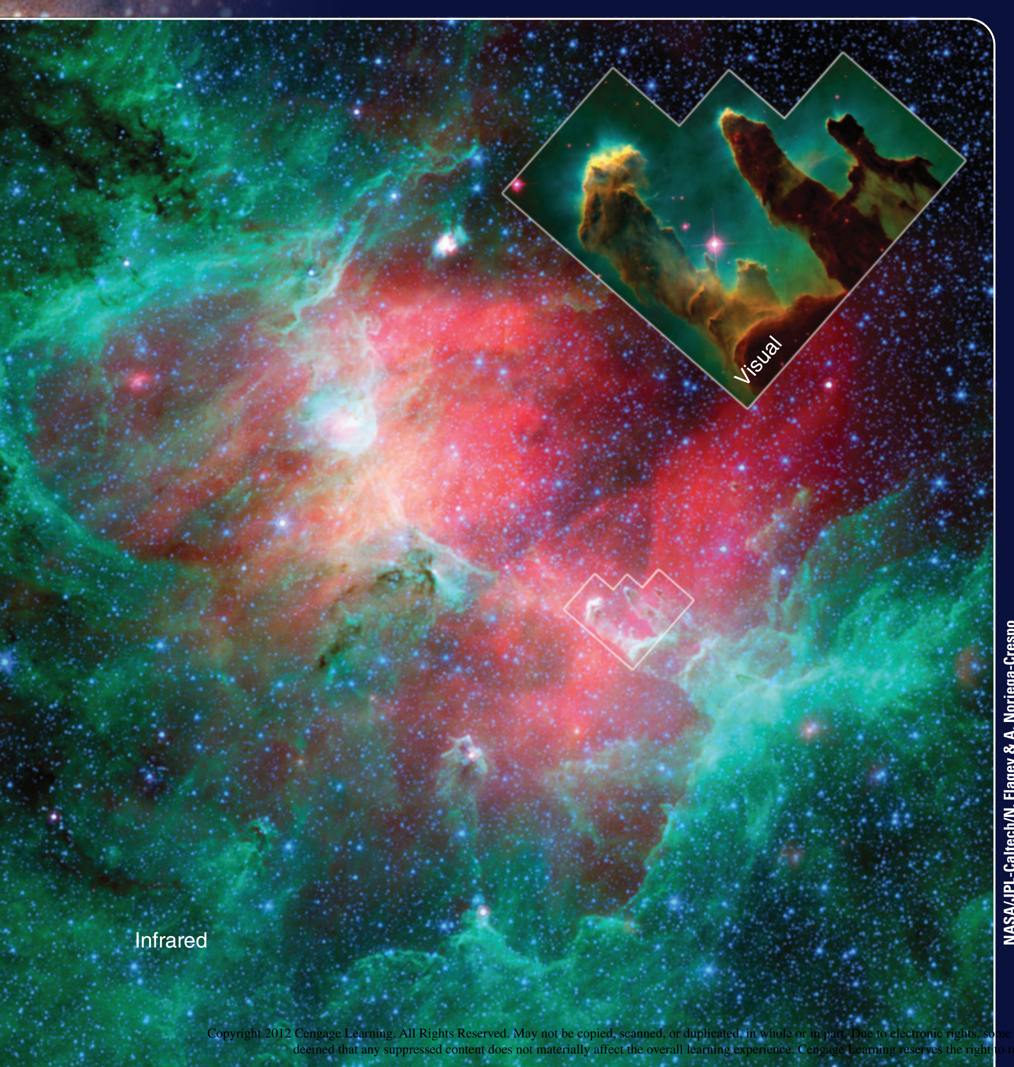


4 At the center of this image, a newborn star is emitting powerful jets to left and right. Where the jets strike the interstellar medium, they produce **Herbig–Haro objects.** The inset shows how irregular the jet is. Such jets can be over a light-year long and contain gas traveling at 100 km/s or more.

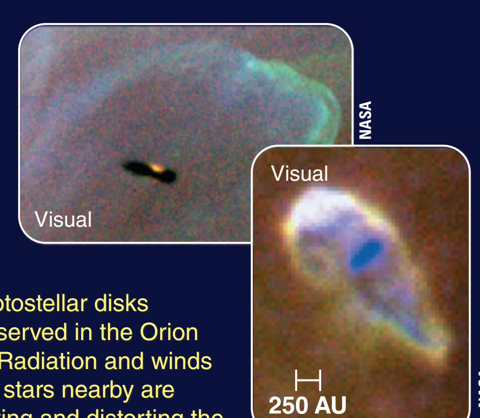
4a Herbig–Haro objects, named after the two astronomers who first described them, are small nebulae that fluctuate in brightness. They appear to be produced by flickering jets from newborn stars exciting the interstellar medium.



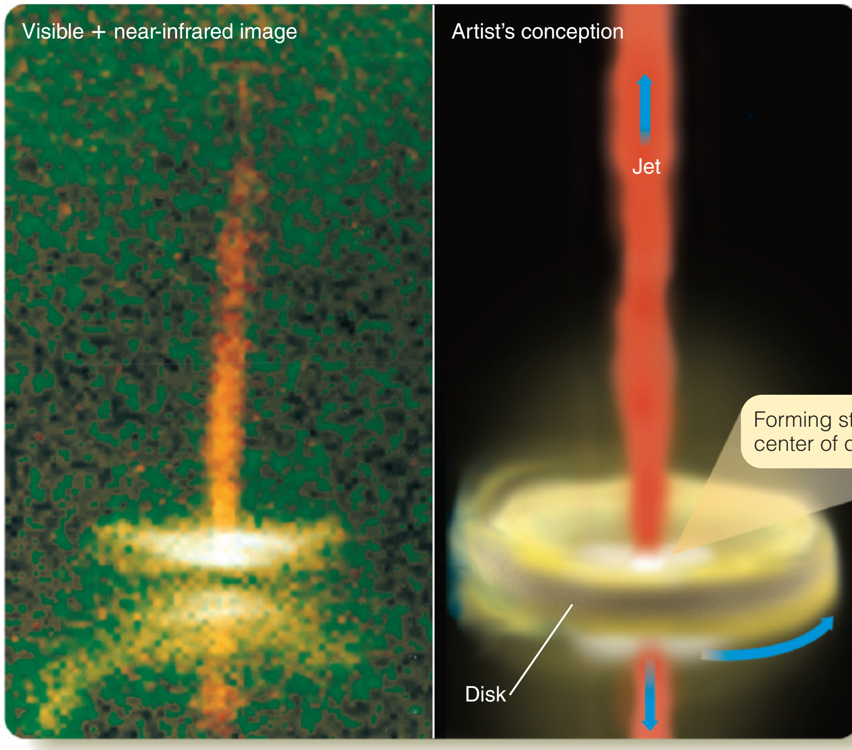
4b Matter flowing into a protostar swirls through a thick disk and, by a process thought to involve magnetic fields, ejects high-energy jets in opposite directions. Observation of these **bipolar flows** is evidence that protostars are surrounded by disks because only disks could focus the flows into jets.



4c Radiation and winds from massive stars have shaped this nebula, Messier 16, and a recent supernova has heated some of the dust (red). Shock waves from the explosion will destroy the Eagle Nebula star-formation pillars (inset) within about 1000 years. Erosion of part of the Eagle Nebula has exposed small fingers and globules of denser gas. About 15 percent of those have formed protostars.



4d Protostellar disks observed in the Orion Nebula. Radiation and winds from hot stars nearby are evaporating and distorting the disks. The disk above is seen silhouetted against the nebula, with light from the protostar embedded in the disk's center apparent at its upper edge. Although bigger than the present size of our solar system, such disks are understood to be likely sites of planet formation.



■ **Figure 14-12**

In the object HH 30, a newly formed star lies at the center of a dense disk of dusty gas that is narrow near the star and thicker farther away. Although the star is hidden from you by the edge-on dusty disk, the star illuminates the inner surface of the disk. Interactions between the in-falling material in the disk and the spinning star eject jets of gas along the axis of rotation. The artist's conception on the right is drawn to help interpret what you see in the *Hubble Space Telescope* image on the left.

SCIENTIFIC ARGUMENT

What evidence can you cite that stars in so-called star forming regions are actually young?

This is a very good example of building a scientific argument because you must cite different forms of evidence to confirm a hypothesis. First, you should note that some extremely luminous stars can't live very long, so when you see such stars, such as the hot blue stars in Orion, you know they must have formed in the last few million years. Many regions of gas and dust contain bright, hot stars that are caught in the act of blowing their nebulae apart; those stars must also have formed recently, or the nebulae would be gone already.

You have other evidence when you look at T Tauri stars. They (1) are usually associated with interstellar gas and dust concentrations, (2) have H–R diagram locations just above the main sequence where you expect stars to be that are still contracting, and (3) often have protostellar disks that are easily destroyed and cannot last long.

There seems to be no doubt that star formation is an ongoing process, and you can revise your argument to discuss how it happens. **What evidence can you cite that protostars are often surrounded by disks of gas and dust?**

14-4 Stellar Structure and Nuclear Fusion

How WOULD YOU GO about making a star? It's simple, really. All you need is something to gather together a star's mass of interstellar medium, maybe a light-year-sized broom, and gravity will do the rest, contracting the cloud first into a protostar and then into a stable star. The structure of a star is surprisingly simple.

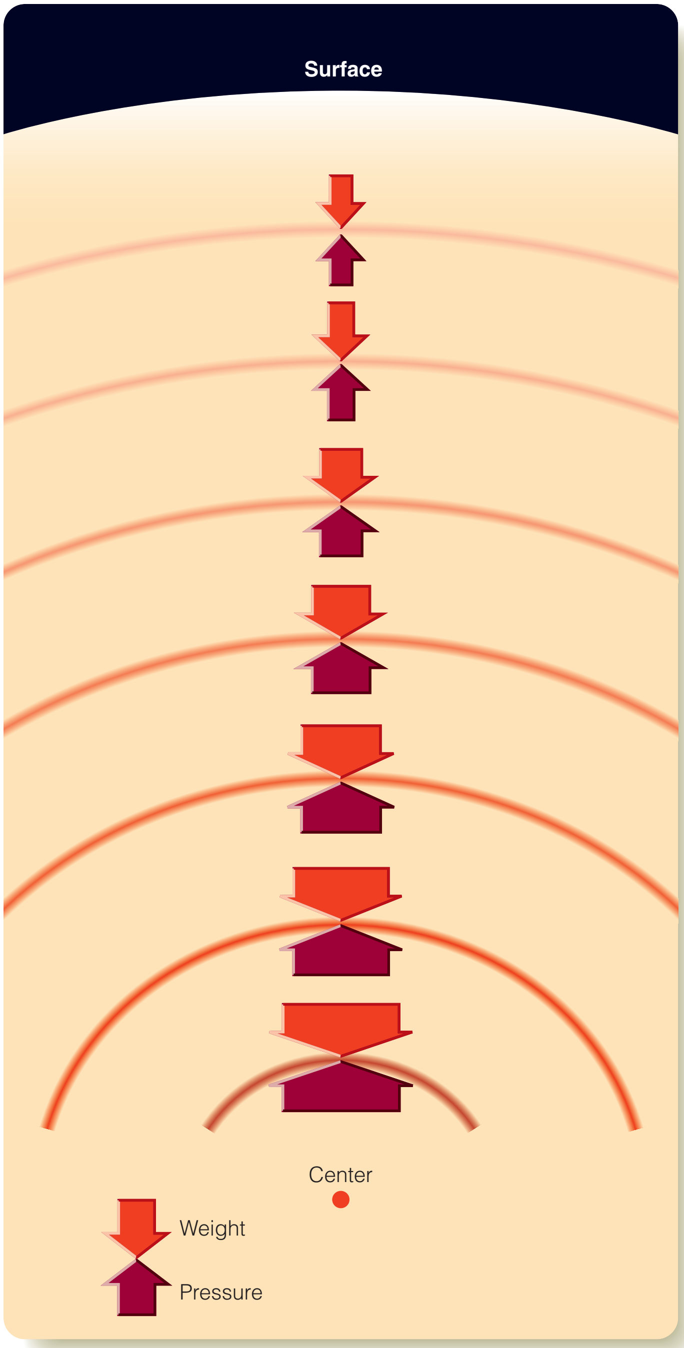
What Keeps a Star Stable?

If there is a single idea in stellar astronomy that can be called crucial, it is the concept of balance. In this section you will discover that stars are held together by their own gravity balanced by the support of their internal heat and pressure. That story will lead your imagination into a region where you yourself can never go—the heart of a star.

Consider the structure of a star, using the basic concept of balance. What is meant here by structure is the variation in properties such as temperature, density, pressure, composition, between the surface of the star and its center. You can think more easily about stellar structure if you imagine that the star is divided into many concentric shells like those in an onion. You can then focus on the temperature, pressure, and density in each shell. These helpful shells exist only in the imagination; stars do not actually have such separable layers.

The weight of each layer in a star must be supported by the layers below (closer to the center). Picture a pyramid of people in a circus stunt—the people in the top row do not have to hold up anybody else; the people in the next row down are holding up the people in the top row, and so on. In a star that is stable, the deeper layers must support the weight of all of the layers above. The inside of a star is made up of gas, so the weight pressing down on a layer must be balanced by gas pressure in that layer. If the pressure is too low, the weight from above will compress and push down the layer; and if the pressure is too high, the layer will expand and lift the layers above.

This stable balance between weight and pressure is called the principle of **hydrostatic equilibrium.** *Hydro* (from the Greek word for water) tells you the material is a fluid, which by definition includes the gases of a star, and *static* tells you the fluid is stable, neither expanding nor contracting. ■ Figure 14-13 shows this hydrostatic balance in the imaginary layers of a star. The



© Cengage Learning 2014

Figure 14-13

The principle of hydrostatic equilibrium says the pressure in each layer of a star must balance the weight on that layer. Consequently, as the weight increases from the surface of a star to its center, the pressure must also increase.

pressure in a gas depends on the temperature and density of the gas. Near the surface there is not much weight pressing down, so the pressure does not need to be high for stability. Deeper in the star, the pressure must be higher, which means that the temperature and density of the gas must also be higher. In other words, the principle of hydrostatic equilibrium tells you that stars must have high temperature, pressure, and density inside to support their own weight and be stable. You can be absolutely sure, from this simple argument, that the pressure inside a star must grow larger with increasing depth to support the weight and keep the star stable, even if you can't directly examine the insides of a star.

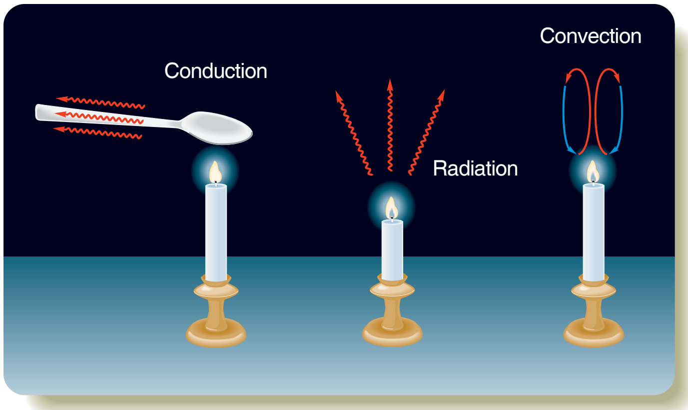
Although the principle of hydrostatic equilibrium can tell you some things about the inner workings of stars, you also need to know how energy flows within stars to completely understand them.

Energy Transport

The surface of a star radiates light and heat into space and would quickly cool if that energy were not replaced. Because the inside of the star is hotter than the surface, energy must flow outward to the surface, where it radiates away. This flow of energy through each shell determines its temperature, which, as you saw earlier, determines how much weight that shell can balance. The movement of energy from the inside to the surface of a star thus plays a crucial role in determining the star's structure.

The law of **energy transport** says that energy must flow from hot regions to cooler regions either by **conduction, convection,** or **radiation.** Conduction is the most familiar form of heat flow. If you hold the bowl of a spoon in a candle flame, the handle of the spoon grows warmer as heat, in the form of motion among the atoms of the spoon, is conducted from atom to atom up the handle (■ Figure 14-14).

The transport of energy by radiation is another familiar experience. Put your hand near a candle flame, and you can feel the heat. What you actually feel are infrared photons—packets of energy—radiated by the flame and absorbed by your hand. Radiation is the principal means of energy transport in the interiors of most stars. Photons are absorbed and reemitted in random directions over and over as energy works its way from the hot interior toward the cooler surface. The flow of energy by radiation is controlled by the **opacity** of the gas, its resistance to movement of



© Cengage Learning 2014

Figure 14-14

The three modes by which energy may be transported from the flame of a candle, as shown here, are the three modes of energy transport within a star. Radiation and convection are much more important than conduction in the interiors of main sequence stars.

radiation. Opacity in turn depends strongly on the temperature: A hot gas is more transparent, less opaque, than a cool gas.

If the opacity is high, radiation cannot flow through the gas easily, and it backs up like water behind a dam. When enough heat builds up, the gas begins to churn as hot gas rises upward and cool gas sinks downward. This heat-driven circulation of a fluid is called convection, the third way energy can move inside a star. You are familiar with convection: The rising wisp of smoke above a candle flame is carried by convection. Energy is carried upward in these convection currents as rising hot gas (red in the right-hand diagram within Figure 14-14) and also as sinking cool gas (blue in the diagram). You learned about convection in the outermost layer of the sun as the cause of visible granulation features, and also as contributing to the solar magnetic activity cycle. Convection in stars is important not only because it carries energy but also because it mixes the gas.

Radiation or convection are the most efficient means of heat movement in the interiors of normal stars; conduction is important only in rare types of stars with extremely high densities.

Nuclear Fusion

There is a **Common Misconception** that "the sun's core is hot because nuclear reactions occur there." That statement reverses cause and effect. Rather, it is correct to say the sun and other stars have nuclear reactions in their cores because the temperature is hot enough there. You observe the sun to be stable; therefore, it must have a certain temperature, pressure, and density in its core. If somehow you could magically turn off the nuclear reactions in its core, then without a source of energy to replace the luminosity pouring out into space the sun would gradually lose its internal energy and start slowly contracting again, as was true during its protostellar phase before the nuclear reactions started.

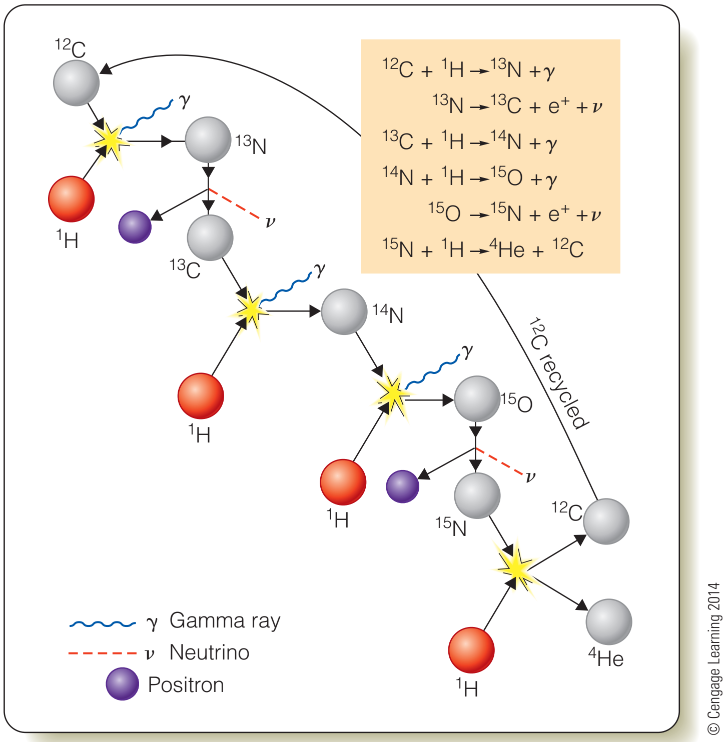
In Chapter 7 you imagined a visit to the center of the sun and discovered that it creates energy through hydrogen fusion in a series of nuclear reactions called the proton-proton chain (look back to Figure 7-12). That reaction begins with the fusion of two protons, which have positive charges and repel each other electrically. You also learned that high-speed collisions are required to overcome what is called the *Coulomb barrier,* the resistance of protons to being combined. High speed for atoms and subatomic particles means high temperature. Consequently, the proton-proton chain cannot effectively occur if the gas temperature is lower than about 4 million Kelvin. The gas density must also be high so the number of collisions per second is large, because a huge total production of energy is needed to keep the sun and other stars stable. Model calculations show that only about the innermost 30 percent of the sun's radius has conditions that support fusion. The rest of the sun, farther from the center, isn't hot and dense enough.

The CNO Cycle

You might expect other stars to fuse hydrogen the same way the sun does, and you would be right for most stars. Some stars, however, fuse hydrogen by another process, and that makes a big difference to their structures.

Look carefully at ■ Figure 14-15 and study the steps in the **CNO cycle.** During the process, four protons combine to make a helium nucleus. In other words, the CNO cycle is really hydrogen fusion and has the same outcome as the proton–proton chain, but it is different in important ways. Notice that carbon-12 begins the cycle and ends the cycle, so the carbon-12 nucleus is recycled over and over. Notice also that the CNO cycle includes steps in which protons (hydrogen nuclei) combine with carbon and nitrogen nuclei. Because carbon and nitrogen nuclei have positive charges respectively six and seven times higher than a proton, the Coulomb barrier is higher for this reaction than for the proton-proton chain. Temperatures higher than 16 million Kelvin are required for significant numbers of protons to combine with carbon and nitrogen nuclei. The temperature at the very center of the sun is estimated to be 15.7 million K, which is why the sun makes nearly all of its energy from the proton-proton chain and only a little (about 10 percent) from the CNO cycle. The CNO cycle produces most of the energy in the hotter centers of main sequence stars more massive than the sun.

The stars in the evening sky look much the same, but you have discovered that they are a diverse group, both outside and inside. Different kinds of stars make their energy in different ways and must also have different internal structures, determined



■ **Figure 14-15**

The CNO cycle uses ^{12}C as a catalyst to combine four hydrogen nuclei (^{1}H) to make one helium nucleus (^{4}He) plus energy. The carbon nucleus reappears at the end of the process, ready to start the cycle over.

by the principle of balance. But, can a star "lose" its balance? How do stars maintain their stability?

The Pressure-Temperature Thermostat

Newborn stars contract and heat up until nuclear fusion begins. The energy flowing outward from the core heats the layers of gas, raises the pressure, and stops the contraction. This leads to an interesting question: How does the star manage to make just enough energy to stop contracting but not to start expanding? The key is the relationship between pressure and temperature that acts like a thermostat to keep the star burning steadily.

Consider what would happen if the reactions begin to produce too much energy. Normally, the nuclear reactions generate just enough energy to balance the inward pull of gravity. If the star makes slightly too much energy, the extra energy flowing out of the star would force its layers to expand slightly, lowering the central temperature and density and slowing the nuclear reactions until the star regained stability. Thus, a star has a built-in regulator, the **pressure-temperature thermostat,** that keeps its nuclear reactions from occurring too rapidly.

The same thermostat keeps the reactions from dying down. Suppose the nuclear reactions begin making too little energy. Then the star would contract slightly, increasing the central temperature and density, which would in turn increase the nuclear energy generation until the star regained stability.

The stability of a star depends on this relation between pressure and temperature. If an increase or decrease in temperature produces a corresponding change in pressure, the thermostat functions correctly, and the star is stable. You will discover in a later section of this chapter how the thermostat accounts for the relationship between mass and luminosity of main-sequence stars. In Chapter 15, you will see what happens to a star when the thermostat breaks down completely and the nuclear fires rage unregulated.

Stellar Models

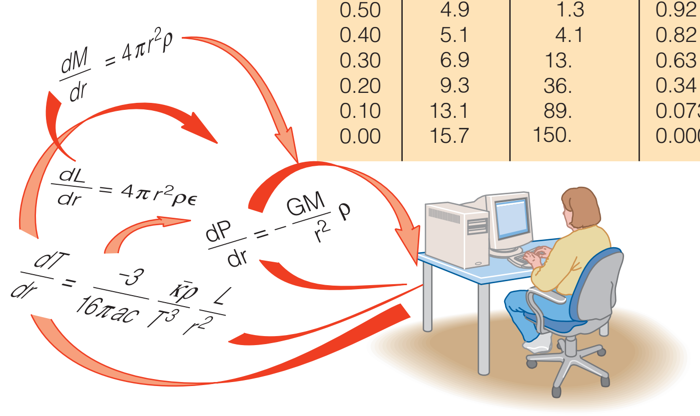
The laws of stellar structure, described in general terms in the previous sections, can be written as mathematical equations. By solving those equations in a special way, astronomers can build a mathematical model of the inside of a star. If you wanted to build a model of a star, you would have to divide the star into about 100 concentric shells and then write down the four equations of stellar structure for each shell. You would then have 400 equations that would have 400 unknowns, namely, the temperature, density, mass, and energy flow in each shell. Solving 400 equations simultaneously is not easy, and the first such solutions, done by hand before the invention of the electronic computer, took months of work. Now a properly programmed computer can solve the equations in a few seconds and print a table of numbers that represents the conditions in each shell of the star. Such a table is a **stellar model**—a mathematical description of the inside of a star (**How Do We Know? 14-2**).

The table shown in ■ Figure 14-16 is a model of the sun. The bottom line, for radius equal to 0.00, represents the center

■ Figure 14-16

A stellar model is a table of numbers that represent conditions inside a star. Such tables can be computed using four laws of stellar structure, shown here in equation form, including the principle of hydrostatic equilibrium and the law of energy transport. The table in this figure describes the present-day sun.

$R/R_\odot$	T (10^6 K)	Density (g/cm^3)	$M/M_\odot$	$L/L_\odot$
1.00	0.006	0.00	1.00	1.00
0.90	0.60	0.009	0.999	1.00
0.80	1.2	0.035	0.996	1.00
0.70	2.3	0.12	0.990	1.00
0.60	3.1	0.40	0.97	1.00
0.50	4.9	1.3	0.92	1.00
0.40	5.1	4.1	0.82	1.00
0.30	6.9	13.	0.63	0.99
0.20	9.3	36.	0.34	0.91
0.10	13.1	89.	0.073	0.40
0.00	15.7	150.	0.000	0.00



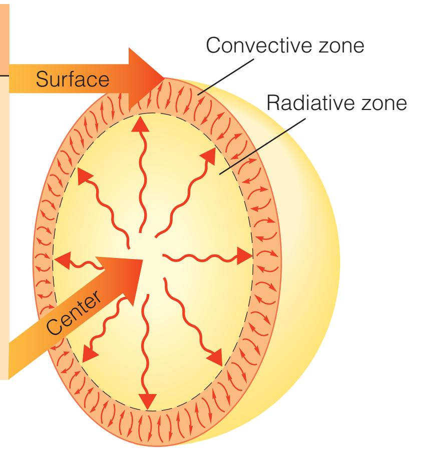


Illustration design by Michael A. Seeds

How Do We Know? | 14-2

Mathematical Models

How can scientists study aspects of nature that cannot be observed directly? One of the most powerful tools in science is the mathematical model, a group of equations or a computer program carefully designed to mimic the behavior of objects and processes that scientists want to study. Astronomers build mathematical models of stars to study the structure hidden deep inside them. Models can speed up the slow evolution of stars and slow down the rapid processes that generate energy. Stellar models are based on only four equations, but other models are much more complicated and may require many more equations.

For example, scientists and engineers designing a new airplane don't just cross their fingers, build it, and ask a test pilot to try it out. Long before any metal parts are made, mathematical models are created to test whether the wing design will generate enough lift, whether the fuselage can support the strain, and whether the rudder and ailerons can safely control the plane during takeoff, flight, and landing. Those mathematical models are put through all kinds of tests: Can a pilot fly with one engine shut down? Can the pilot recover from sudden turbulence? Can the pilot land in a crosswind? By the time the test pilot rolls the plane down the runway for the first time, the mathematical models have flown many thousands of miles.

Scientific models are only as good as the assumptions that go into them and must be compared with the real world at every opportunity. If you are an engineer designing a new airplane, you can test your mathematical models by making measurements in a wind tunnel. Models of stars are much harder to test against reality, but they do predict some observable things. Stellar models predict the existence of a main sequence, the mass-luminosity relation, the observed numbers of giant and supergiant stars, and the shapes of H–R diagrams. Without mathematical models, astronomers would know little about the lives of the stars, and designing new airplanes would be a very dangerous business.

The Boeing Company

Before any new airplane flies, engineers build mathematical models to test its stability.

of the sun, and the top line, for radius equal to 1.00, represents the surface. The other lines in the table tell you the temperature and density in each shell, the mass inside each shell, and the fraction of the sun's luminosity flowing outward through the shell. You can use the table to study conditions in the sun. For example, the bottom line tells you the temperature at the center of the sun is over 15 million Kelvin. At such a high temperature the gas is highly transparent, and energy flows as radiation. Nearer the surface, the temperature is lower, the gas is more opaque, and energy is carried by convection.

Although this may sound simple, it is actually a highly challenging problem involving nuclear and atomic physics, thermodynamics, and sophisticated mathematical methods. Computers make the rapid calculation of detailed stellar models possible, and astronomy has advanced tremendously by comparing such models with observations to study the structure and evolution of stars. For example, the summary of star formation in this chapter is based on thousands of stellar models. You will continue to rely on theoretical models as you study the lives and deaths of stars in the next chapter.

14-5 Main-Sequence Stars

When a contracting protostar begins to fuse hydrogen, it stops contracting and becomes a stable main-sequence star. The most massive stars are so hot that they light up the remaining nearby gas in beautiful nebulae, as if announcing their birth (Figure 14-5, Figure 14-9). After the gas of the nebula disperses, the star begins its long, mostly uneventful life as a main-sequence star. In a careful census of true stars (excluding white dwarfs), about 90 percent are main sequence stars.

The Relation Among Stellar Mass, Luminosity, and Lifetime

You learned in Chapter 13 that observations of the temperatures and luminosities of stars show that main-sequence stars obey a simple rule—the more massive a star is, the more luminous it is. That rule, the mass-luminosity relation, is the consequence of the mechanisms causing the stability of main-sequence stars. In fact, the mass-luminosity relation is predicted by theories of stellar structure, giving astronomers direct observational confirmation of those theories.

To understand the mass-luminosity relation, recall (1) the principle of hydrostatic equilibrium, which says that pressure balances weight, and (2) the pressure-temperature thermostat, which regulates energy production. A star that is more massive than the sun has more weight pressing down on its interior, so the interior must have a higher pressure to balance that weight. That means the massive star's automatic pressure-temperature thermostat must keep the gas in its interior hot and the pressure high. A star less massive than the sun has less weight on its interior and thus needs less internal pressure; therefore, its pressure-temperature thermostat is set lower. To sum up, massive stars are more luminous because they must make more energy to support more weight.

The mass-luminosity relation tells you why the main sequence must have a lower end. Low-mass objects have lower pressures and temperatures in their cores, and those with less than about 0.08 solar mass are too cool to allow hydrogen fusion. Called **brown dwarfs,** such objects have diameters only about the size of Jupiter (ten times the size of Earth), and although they are still warm from contraction, they do not generate energy. They have contracted as far as they can and are slowly cooling off. Brown dwarfs fall in the gap between low-mass M stars and massive planets like Jupiter. You met them in Chapter 13 (page 266) in the form of new stellar spectral types cooler than M; the warmer brown dwarfs fall in spectral class L and the cooler in spectral class T. Brown dwarfs would look dull reddish-orange to your eyes (hence the term *brown*), but they emit most of their energy in the infrared. Brown dwarfs are so cool that liquid and solid particles of silicates, metals, and other minerals can condense to form cloud layers in their atmospheres. Evidently, brown dwarfs have rocks for weather.

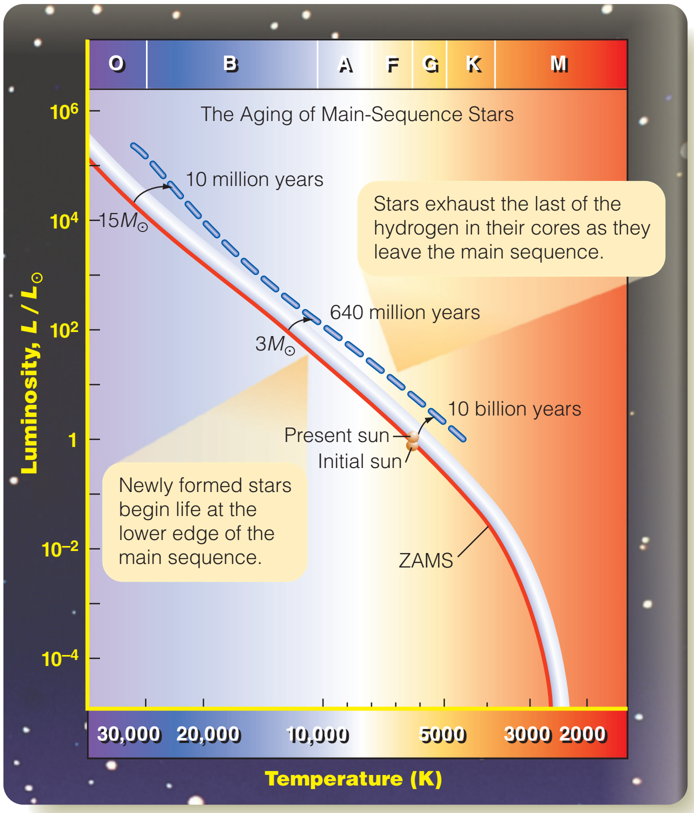
Because they are so small and cool, brown dwarfs are very low-luminosity objects and are difficult to find. Nevertheless, hundreds are known. The evidence shows that nature does indeed make brown dwarfs when a protostar does not have enough mass to begin hydrogen fusion. This observational detection of the lower end of the main sequence further confirms the theories of stellar structure. Now that you know how stars form and how they maintain their stability through the mass-luminosity relation, you can predict the lives and deaths of main-sequence stars.

The Life of a Main-Sequence Star

A star on the main sequence is stable, with its gravity balanced by outflowing energy from fusing hydrogen. The balance must shift as the star consumes its fuel, so that even apparently stable main-sequence stars have interiors that are slowly changing.

As a main-sequence star fuses its hydrogen, four hydrogen nuclei are converted into single helium nuclei, and the total number of particles in its interior decreases. Each newly made helium nucleus exerts the same pressure as one hydrogen nucleus, but because the gas has fewer nuclei its total pressure is less. This unbalances the gravity–pressure stability, and gravity squeezes the core of the star more tightly. As the core slowly contracts, its temperature increases, and the nuclear reactions run faster, releasing more energy. This additional energy flowing outward forces the outer layers to expand. As the star becomes gradually larger, it becomes more luminous, and eventually the expansion begins to cool the surface.

As a result of these gradual changes in main-sequence stars, the main sequence is not a sharp line across the H–R diagram but rather a band. Stars begin their stable lives fusing hydrogen on the lower edge of this band, which is known as the **zero-age main sequence (ZAMS),** but gradual changes in luminosity and surface temperature move the stars upward and slightly to the right, as shown in ■ Figure 14-17. By the time they reach the upper edge of the main sequence, they have exhausted nearly all



■ **Figure 14-17**

A newborn star reaches stability with properties that place it on the H–R diagram at the lower edge of the main sequence band, along the line called the zero-age main sequence (ZAMS). As a star converts hydrogen in its core into helium, it moves slowly away from the ZAMS and across the main sequence, becoming slightly more luminous and cooler. Once a star consumes all of the hydrogen in its core, it can no longer remain a stable main-sequence star. More massive stars age rapidly, but less massive stars use up the hydrogen in their cores more slowly and live longer main-sequence lives.

the hydrogen in their centers. Thus you find main-sequence stars scattered throughout the band at various stages of their main-sequence lives.

The sun is a typical main-sequence star; as it undergoes these gradual changes, Earth will suffer. When the sun began its main sequence life almost 5 billion years ago, it was only about 70 percent as luminous as it is now, and by the time it leaves the main sequence in another 5 billion years, the sun will have twice its present luminosity. Long before that, the rising luminosity of the sun will drastically modify Earth's climate, and ultimately boil our oceans away. Life on Earth will probably not survive these changes in the sun, but we have a billion years or more to prepare.

A normal star spends 90 percent of its life on the main sequence. That explains why 90 percent of all true stars are main-sequence stars—you are most likely to see a star during that long, stable period while it is on the main sequence. To illustrate, imagine that you photograph a crowd of 20,000 people. Everyone sneezes now and then, but the act of sneezing is very short compared with a human lifetime, so you would expect to find that very few people in your photograph are caught in the act of sneezing. Rather, most people in the photo would be in the much more common nonsneezing state. Your view of the universe is like that snapshot; you see most stars on the main sequence, where they spend most of their time.

The number of years a star spends on the main sequence depends on its mass (■Table 14-1). Massive stars consume fuel rapidly and live short lives, but low-mass stars conserve their fuel and shine for billions of years. For example, a 25-solar-mass star will exhaust its hydrogen and die in only about 7 million years. The sun has enough fuel to last about 10 billion years. The red dwarfs, although they have little fuel, use it up very slowly and may be able to survive for 100 billion years or more. **Reasoning with Numbers 14-1** explains how you can quickly estimate the life expectancies of stars from their masses.

■ Table 14-1 | Main-Sequence Stars: Masses, Luminosities and Lifetimes

Spectral Type	Mass (Sun = 1)	Luminosity (Sun = 1)	Years on Main Sequence
O5	40	400,000	1×10^6
B0	15	13,000	11×10^6
A0	3.5	80	440×10^6
F0	1.7	6.4	3×10^9
G0	1.1	1.4	8×10^9
K0	0.8	0.46	17×10^9
M0	0.5	0.08	56×10^9

Reasoning with Numbers | 14-1

The Life Expectancies of Stars

You can estimate the amount of time a star spends on the main sequence—its life expectancy, τ—by estimating the amount of fuel it has and dividing by the rate at which it consumes that fuel:

$$\tau = \frac{\text{fuel supply}}{\text{rate of consumption}}$$

The amount of fuel a star has is proportional to its mass M, and the rate at which it uses up its fuel is proportional to its luminosity L. Thus its life expectancy must be proportional to M/L. You can simplify this equation further because, as you saw in Chapter 13, the luminosity of a star depends on its mass raised to the 3.5 power ($L = M^{3.5}$). So the life expectancy is

$$\tau = \frac{M}{M^{3.5}}$$

which is the same as:

$$\tau = \frac{1}{M^{2.5}}$$

If you express the mass in solar masses, the lifetime will be in solar lifetimes.

Example: How long can a 4-solar-mass star live? **Solution:**

$$\tau = \frac{1}{4^{2.5}}$$

$$= \frac{1}{32} \text{ solar lifetimes}$$

Solar models show that the sun, presently 5 billion years old, will last another 5 billion years. Thus a solar lifetime is approximately 10 billion years, and a 4-solar-mass star will last for about

$$\tau = \frac{1}{32} \times (10 \times 10^9 \text{ yr})$$

$$= 310 \times 10^6 \text{ years}$$

Nature makes more low-mass stars than high-mass stars, but this fact is not sufficient to explain the vast numbers of low-mass stars that fill the galaxy. An additional factor is stellar lifetime. Because low-mass stars live long lives, there are more of them in the sky than massive stars. Look at page 283 and notice how much more common the lower-main-sequence stars are than the massive O and B stars. The main-sequence K and M stars are so faint they are difficult to locate, but they are very common. The main sequence O and B stars are luminous and easy to locate; but because of their fleeting lives there are never more than a few visible in the sky.

SCIENTIFIC ARGUMENT

What would happen if the sun stopped generating energy?

Sometimes one of the best ways to test your understanding is to build an argument based on an altered situation. Stars are supported by the outward flow of energy generated by nuclear fusion in their interiors. That energy keeps each layer of the star just hot enough for the gas pressure to support the weight of the layers above. Each layer in the star must be in hydrostatic equilibrium; that is, the inward weight must be balanced by outward pressure. If the sun stopped making energy in its interior, nothing would happen at first, but over many thousands of years the loss of energy from its surface would reduce the sun's ability to hold steady against its own gravity, and it would begin to contract. You wouldn't notice much for 100,000 years or so, but eventually the sun would lose its battle with gravity.

Stars are elegant in their simplicity—nothing more than a cloud of gas held together by gravity and warmed from the inside by nuclear fusion. Now build a different argument. **How does the star manage to make exactly the right amount of energy to support its weight?**

What Are We?
Myth-Makers and Explainers

On cold winter nights when the sky is clear and the stars are bright, Jack Frost paints icy lacework across your windowpane. That's a fairy tale, of course, but it is a graceful evocation of the origin of frost. We humans are explainers, and one way to explain the world around us is to create myths.

An ancient Aztec myth tells the story of the origin of the moon and stars. The stars, known as the Four Hundred Southerners, and the moon, the goddess Coyolxauhqui, plotted to murder their unborn brother, the great war god Huitzilopochtli. Hearing their plotting, he leaped from the womb fully armed, hacked Coyolxauhqui into pieces, and chased the stars away. You can see the Four Hundred Southerners scattered across the sky, and each month you can see the moon chopped into pieces as it passes through its phases.

Stories like these explain the origins of things and can make our universe seem more understandable. Science is a natural extension of our need to explain the world. The stories have become sophisticated scientific theories and are tested over and over against reality, but we humans build those theories for the same reason people used to tell myths.

Study and Review

Summary

- The **interstellar medium (ISM) (p. 289),** the gas and dust between the stars, is mostly concentrated near the plane of our Milky Way Galaxy and has an average density of about one atom per cubic centimeter.
- **Interstellar extinction (p. 289),** or dimming, makes the distant stars look fainter than they should. **Interstellar reddening (p. 289)** makes distant stars appear too red because dust particles in the interstellar medium scatter blue light more easily than red light. The dependence of this extinction on wavelength shows that the scattering dust particles are very small. The dust is made of carbon, silicates, iron, ice, and other cosmically abundant substances.
- The interstellar gas is cold and has a very low density, and this makes **interstellar absorption lines (p. 290)** much narrower than the spectral lines produced in stars. Such lines are usually obvious in stellar spectra because they represent ions that cannot exist in the atmospheres of the stars.
- The low-density gas of the interstellar medium also produces **interstellar emission lines (p. 291)** at many wavelengths. The **21-cm radio line (p. 291)** is an emission line produced when the electrons in hydrogen atoms change the direction of their spin and emit radio-wavelength photons. This radiation allows radio astronomers to map the distribution of neutral hydrogen gas in the interstellar medium.
- About 70 percent of the mass of the interstellar medium is hydrogen gas, and 28 percent is helium. About 2 percent is atoms heavier than helium. This is approximately the same composition as the sun and other stars.
- **Interstellar dust (p. 290)** makes up roughly 1 percent of the mass of the interstellar medium. The remaining 99 percent of the mass is gas.
- Large, dense clouds of gas and dust are called **molecular clouds (p. 290)** because they are so dense that molecules can form inside them. The largest of these are called **giant molecular clouds (p. 296)** and are the sites of star formation. Radio, infrared, and X-ray telescopes have detected emission from over 150 different molecules in the interstellar medium.
- A **nebula (p. 292)** is a cloud of gas in space, and an **HII region (p. 294),** also known as an **emission nebula (p. 294),** is produced when ultraviolet radiation from hot stars ionizes nearby gas, making it glow like a giant neon sign. The red, blue, and violet Balmer lines blend together to produce the characteristic pink-red color of ionized hydrogen.
- A **reflection nebula (p. 294)** is produced by gas and dust illuminated by a star that is not hot enough to ionize the gas. Rather, the dust scatters the starlight to produce a reflection of the stellar absorption spectrum. Because shorter-wavelength photons scatter more easily than longer-wavelength photons, reflection nebulae look blue. The daytime sky looks blue for the same reason.
- A **dark nebula (p. 295)** is a cloud of gas and dust that is noticeable because it blocks the light of distant stars. The irregular shapes of these dark nebulae reveal the turbulence in the interstellar medium.
- The existence of massive hot stars that cannot live very long, such as the prominent stars of Orion, is strong evidence that stars have formed recently.
- The gravity of giant molecular clouds makes them contract, but that is resisted by thermal energy in the gas, and also by magnetic fields, rotation, and turbulence. In at least some cases, clouds are compressed by passing shock waves, also known as **shocks (p. 293),** and star formation is triggered. The birth of massive stars can produce shock waves that trigger further star formation.
- **Bok globules (p. 300)** are small dark nebulae, some of which may be contracting to form stars.
- The cold gas of interstellar space heats up as it contracts because the atoms fall inward in **free-fall collapse (p. 296)** and pick up speed. When the atoms collide, gravitational energy is converted into thermal energy.
- Collapsing clouds dense enough to be opaque slow their contraction and become **protostars (p. 297).** Protostars form deep inside gas and dust **cocoons (p. 297)** and are not directly visible at visual wavelengths until the cocoons dissipate.
- The visible Orion Nebula is only a small part of a much larger dusty molecular cloud. Ionization by ultraviolet photons from the hottest star is ionizing the gas, lighting up the nebula and making it glow brightly. Infrared observations reveal clear evidence of active star formation deeper in the molecular cloud just to the northwest of the Trapezium.
- Protostars become visible as they cross the **birth line (p. 302)** in the H–R diagram. **T Tauri (p. 304)** stars are solar-mass objects that have just emerged from their cocoons and are located in the H–R diagram between the birth line and the main sequence. Stars that are nearing the main sequence stage are generally termed **Young Stellar Objects (YSOs) (p. 302).**
- Many, perhaps most, protostars form surrounded by dusty **protostellar disks (p. 298),** and jets of gas can be emitted as **bipolar flows (p. 299)** along the axis of the spinning disk. Where the jets push into the surrounding gas, they can form nebulae called **Herbig-Haro objects (p. 299).**
- **Associations (p. 303),** including **T associations (p. 303)** and **OB associations (p. 303),** are groups of stars born together but not bound to each other by their mutual gravity. The presence of these associations in an area is evidence of recent star formation.
- The principle of **hydrostatic equilibrium (p. 306)** says that the weight pressing down on a layer of gas in a star must be balanced by the pressure in the gas for the star to be stable. That shows that the inner layers of stars must be hotter because they must support more weight.
- The law of **energy transport (p. 307)** states that energy must flow from hot regions such as the core of the sun to relatively cool regions such as the surface of the sun by **conduction (p. 307), radiation (p. 307),** or **convection (p. 307).** The **opacity (p. 307)** of a gas is its resistance to the flow of radiation. In regions where the opacity of the gas does not permit radiation to carry away enough energy, the gas can churn in convection. One reason convection is important in stellar evolution is that it can mix material between inner and outer parts of stars. Conduction is less efficient than radiation or convection except in a few rare types of stars with very high densities.
- Many stars make their energy the same way the sun does, using the proton–proton chain, which operates only at temperatures above 4 million K needed to overcome the Coulomb barrier of electrical repulsion between positively charged atomic nuclei. The **CNO cycle (p. 308)** is more efficient, but it requires a higher temperature than the proton–proton chain, at least 16 million K, because of the larger repulsion by carbon and other nuclei. Both processes combine four hydrogen nuclei to make one helium nucleus plus energy.
- The relationship between pressure and temperature, the **pressure–temperature thermostat (p. 309),** ensures that the star generates just enough energy to support itself.
- Astronomers can study the interiors of stars and the way they change over time by calculating detailed **stellar models (p. 309)** based on the four laws above. The mass–luminosity relation among main-sequence stars can be understood from stellar models. More massive stars have more weight to support, so that makes them more luminous.
- The main sequence has a lower end because stars less massive than 0.08 solar mass cannot get hot enough to begin hydrogen fusion. Such objects become **brown dwarfs (p. 311).**
- When a star first begins fusing hydrogen into helium, it is located on the H–R diagram along the **zero-age main sequence (ZAMS) (p. 311).** As hydrogen nuclei are converted to helium nuclei, the total number of nuclei in the star's core declines. As a result the core slowly contracts, and the outer layers gradually expand, making the star gradually move upward and to the right across the band of the main sequence.
- How long a star can remain on the main sequence depends on its mass. The more massive a star is, the faster it uses up its hydrogen fuel. A 25-solar-mass star will exhaust its hydrogen and die in only about 7 million years, but the sun is expected to last for 10 billion years.

Review Questions

1. Why do distant stars look redder than their spectral types suggest?
2. What evidence can you cite that the interstellar medium contains both gas and dust?
3. What evidence visible to human eyes can you cite that the spaces between the stars are not totally empty?
4. How do the spectra of HII regions differ from the spectra of reflection nebulae? Why?
5. How is the blue color of a reflection nebula related to the blue color of the daytime sky?
6. Why are interstellar lines so narrow?
7. What factors resist the contraction of a cloud of interstellar matter?
8. Explain the different ways a giant molecular cloud can be triggered to contract.
9. How does a contracting protostar convert gravitational energy into thermal energy?
10. What evidence is there that (a) star formation has occurred recently? (b) protostars really exist? (c) the Orion region is actively forming stars?
11. How does the geometry of bipolar flows and Herbig-Haro objects support the hypothesis that protostars are surrounded by rotating disks?
12. Describe the principle of hydrostatic equilibrium as it relates to the internal structure of a star.
13. How does the CNO cycle differ from the proton-proton chain? How is it similar?
14. How does the pressure-temperature thermostat control the nuclear reactions inside stars?
15. Step-by-step, explain how energy flows from the center of the sun to Earth.
16. Why is there a mass–luminosity relation?
17. Why is there a lower limit to the mass of a main-sequence star?
18. Why does a star's life expectancy depend on its mass?
19. **How Do We Know?** Why are scientists free to adjust their hypotheses but not their facts?
20. **How Do We Know?** How can mathematical models help you understand natural processes that occur in locations or with time scales that make them impossible to observe directly?

Discussion Questions

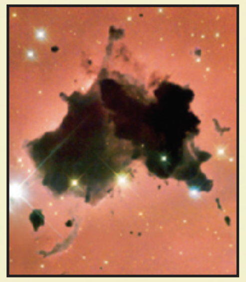
1. When you see distant streetlights through smog, they look dimmer and redder than they do normally. But when you see the same streetlights through fog or falling snow, they look dimmer but not redder. Use your knowledge of the interstellar medium to discuss the relative sizes of the particles in smog, fog, and snowstorms compared to the wavelength of light.
2. If you could see a few stars through a dark nebula, how would you expect their spectra and colors to differ from similar stars just in front of the dark nebula?
3. Ancient astronomers, philosophers, and poets assumed that the stars were eternal and unchanging. Is there any observation they could have made or any line of reasoning that could have led them to conclude that stars don't live forever?
4. How does hydrostatic equilibrium relate to hot-air ballooning?

Problems

1. Extinction dims starlight by about 1 magnitude per 1000 pc. What fraction of photons survives a trip of 1000 pc? (*Hint:* Consider the definition of the magnitude scale in Chapter 2.)
2. The density of air in a child's balloon 20 cm in diameter is roughly the same as the density of air at sea level, 10^{19} particles/cm^3. To how large a diameter would you have to expand the balloon to make the gas inside the same density as the interstellar medium, about 1 particle/cm^3? (*Note:* The volume of a sphere is $\frac{4}{3}\pi r^3$.)
3. The dust in a molecular cloud has a temperature of about 50 K. At what wavelength does it emit the maximum energy? (*Hint:* Use Wien's law, Chapter 6.)
4. A giant molecular cloud is 30 pc in diameter and has a density of 1000 hydrogen atoms/cm^3. What is its total mass in kilograms? (*Note:* The volume of a sphere is $\frac{4}{3}\pi r^3$, and the mass of a hydrogen atom is 1.67×10^{-27} kg.)
5. The expanding bubble of hot gas inflated by the cluster of new stars in its center, shown in Figure 14-7a, has a diameter of about 70 ly. If the bubble is 170,000 ly from Earth, what is the observed diameter of the bubble in arc seconds? (*Hint:* Use the small-angle formula, Chapter 3.)
6. If a giant molecular cloud is 50 pc in diameter and a shock wave can sweep through it in 2 million years, how fast is the shock wave going in kilometers per second?
7. If a giant molecular cloud has a mass of 10^{35} kg and it converts 1 percent of its mass into stars during a single encounter with a shock wave, how many stars can it make? Assume the stars each contain 1 solar mass.
8. The hottest star in the Orion Nebula has a surface temperature of 40,000 K. At what wavelength does it radiate the most energy? (*Hint:* See Wien's law, Chapter 6.)
9. An O star with a lifetime of 0.5 million years drifts away from its birthplace in the center of a molecular cloud at a speed of 10 km/s. If the cloud has a diameter of 10 pc, can the O star escape the cloud while the star is still on the main sequence?
10. The gas in a bipolar flow can travel as fast as 100 km/s. If the length of the jet is 1 ly, how long does it take for a blob of gas to travel from the protostar to the end of the jet?
11. If a protostellar disk is 200 AU in radius and the disk plus the forming star together contain 2 solar masses, what is the orbital speed at the outer edge of the disk in kilometers per second? (*Hint:* See the formula for circular orbital velocity, Chapter 4.)
12. If a contracting protostar is five times the radius of the sun and has a temperature of only 2000 K, how luminous will it be? (*Hint:* See the relation between luminosity, radius, and temperature for stars, Chapter 13.)
13. If a T Tauri star is the same temperature as the sun but is ten times more luminous, what is its radius? (*Hint:* See the relation between luminosity, radius and temperature for stars, Chapter 13.)
14. Circle all of the ^{1}H and ^{4}He nuclei in Figure 14-15 and explain how the CNO cycle can be summarized by $4\ {}^1\text{H} \rightarrow {}^4\text{He} + \text{energy}$.
15. How much energy is produced when the CNO cycle converts 1 kg of mass into energy? (Hint: Is your answer the same, or different, if the mass is fused by the proton–proton chain?)
16. What is the life expectancy of a 16-solar-mass star?

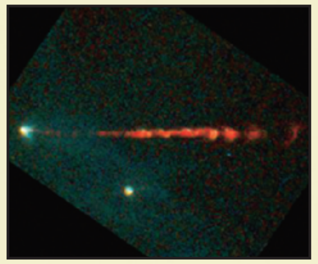
Learning to Look

1. The figure to the right shows a dark globule of dusty gas. What do you think that globule would look like if you could see it from the other side?



NASA and the Hubble Heritage Team, STScI/AURA

2. The star at right appears to be ejecting a jet of gas. What is happening to this star?



STScI and NASA

Great Debates

1. *Jets and Disks.* Jets are seen to emit from the center regions of protostars that are surrounded by dusty disks. Are the jets emitted from the star, the disk, or both?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Research the answers to this question previously suggested by astronomers to support your claim.
   c. Cite your sources.

2. *Is the Sun Ordinary?* If the sun formed from the ISM and from an ordinary protostellar disk, is the sun an ordinary star?
   a. Use the data from your text on average stellar mass, radius, age, number, absolute magnitude or luminosity, temperature or spectral class or color, and metalicity to argue whether the Sun is an ordinary or unusual star.
   b. What's the evidence? Cite table and figure numbers as well as page and paragraph numbers in your answer.
   c. Cite your sources.

3. *Nebula Nickname.* Usually nebulae are nicknamed according to what they look like. What nickname should the nebula in Figure 14-5 on page 293 have? Argue your point by relating features of the that nebula to your nickname.
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

4. *HH 30.* Using the visible plus near-infrared image of HH30 in Figure 14-12, decide whether the disk is accreting material to the protostar. Based on your answer, decide whether the artist's conception needs to be redrawn. Are artist's conceptions scientific models? Should you believe artist's conceptions?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Reddening and Extinction
- Scattering in Earth's Atmosphere
- Star Birth Clouds
- Bipolar Outflow
- Conduction, Convection, and Radiation
- CNO Cycle
- Future of the Sun

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Massive red supergiant star V838 Monocerotis is evolving rapidly. In early 2002, it flared up, and the expanding shell of light from that flare is illuminating surrounding gas that was probably ejected from the star in an earlier outburst. The spikes are produced by diffraction in the telescope.

NASA/Hubble Heritage Project

15

The Deaths of Stars

Guidepost

The preceding chapter described how stars are born and how they resist their own gravity by fusing nuclear fuels in the center and keeping their central pressure high. But they can last only as long as their fuel.

Now you are ready to consider the fate of the stars. Here you will find the answers to five important questions:

- **What happens to a star when it uses up the last of the hydrogen in its core?**
- **What evidence shows that stars really evolve?**
- **How will the sun die?**
- **What happens if an evolving star is in a binary system?**
- **How do massive stars die?**

The deaths of stars are important because life on Earth depends on the sun but also because the deaths of massive stars create the atomic elements of which you are made. If stars didn't die, you would not exist.

In the chapters that follow, you will discover that some of the matter that was once stars becomes trapped in dead ends—white dwarfs, neutron stars, and black holes. But some matter from dying stars escapes back into the interstellar medium and is incorporated into new stars and the planets that circle them. The deaths of stars are part of a great cycle of stellar birth and death that includes our sun, our planet, and us.

Natural laws have no pity.

ROBERT HEINLEIN,
THE NOTEBOOKS OF LAZARUS LONG

GRAVITY IS PATIENT. Stars generate tremendous energy resisting their own gravity, but no star has an infinite supply of fuel for its nuclear reactions. When the fuel runs out, the star dies.

All over the sky, astronomers find beautiful nebulae that were puffed gently into space by dying stars. In addition, astronomers occasionally see a new star appear in the sky, grow brighter, then fade away after a few weeks or a year. You will discover that what looks like a new star in the sky, is either a **nova,** the eruption of a very old dying star, or a **supernova,** the violent explosive death of an aging star. Modern astronomers find a few novae (plural of *nova*) each year, but supernovae (plural) are so rare that there are only one or two each century in our galaxy. Astronomers know that stars die because they occasionally see supernovae flare in other galaxies and because telescopes reveal the remains of stars that have already died (■ Figure 15-1).

The mass of a star is critical in determining its fate. Massive stars can die in violent supernova explosions, but lower-mass stars die quiet deaths. To follow the evolution of stars to their graves, you can start by following the life story of a sunlike, medium-mass star as it becomes a giant star. Then you can see how stars of different masses end their lives.

15-1 Giant Stars

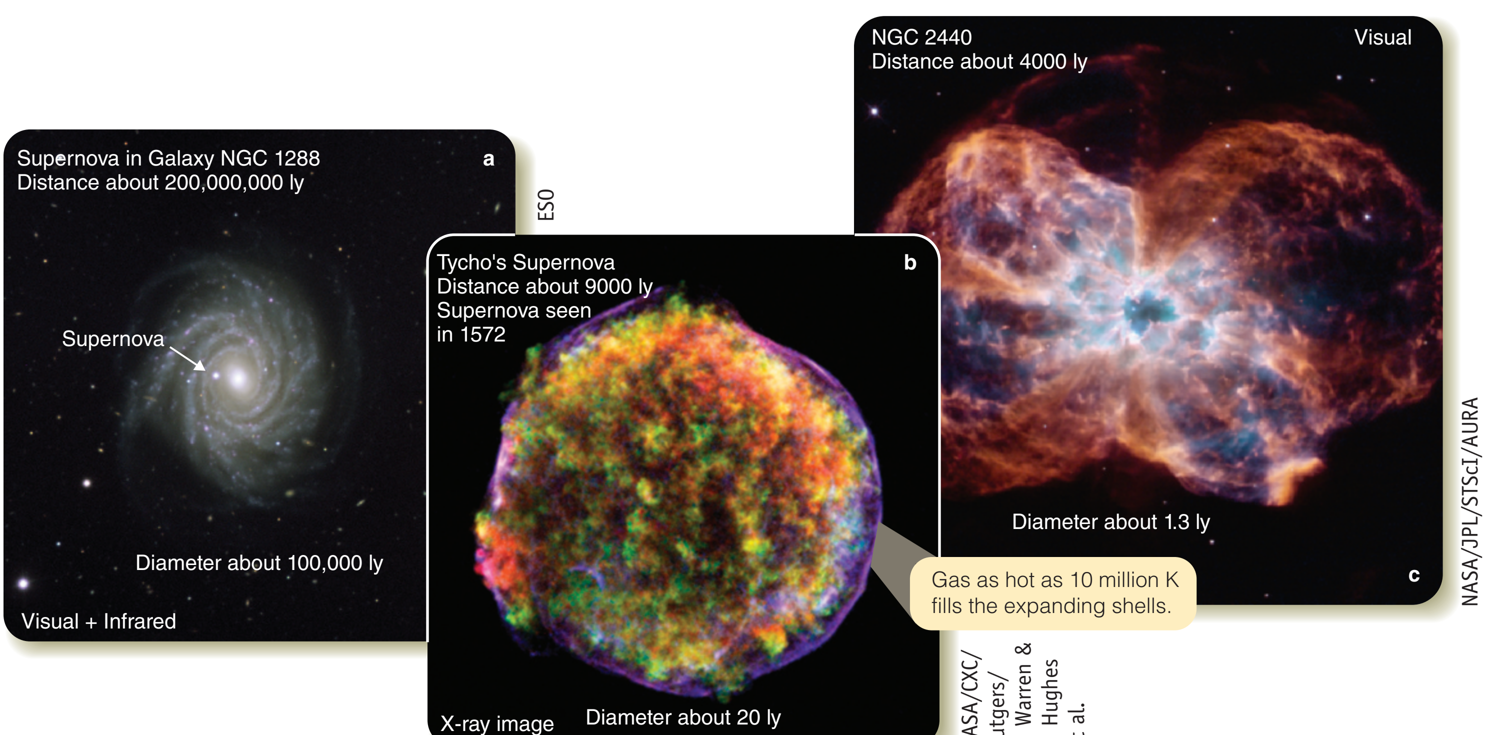
AS YOU LEARNED in the previous chapter, a main-sequence star generates its energy by nuclear fusion reactions that combine hydrogen to make helium. The period during which the star fuses hydrogen lasts a long time, and the star remains on the main sequence for 90 percent of its total existence as an energy-generating star. When the hydrogen is exhausted, however, the star begins to evolve rapidly.

Expansion into a Giant

The nuclear reactions in a main-sequence star's core fuse hydrogen to produce helium. Because the core is cooler than 100,000,000 K, the helium can't overcome the Coulomb barrier to fuse in nuclear reactions, so it accumulates at the star's center like ashes in a fireplace. Initially, this helium ash has little effect on the star, but as hydrogen is exhausted and the stellar core becomes almost pure helium, the star loses the ability to generate the nuclear energy that opposes gravity. As soon as the energy generation starts to die down, gravity begins making the core contract.

■ Figure 15-1

Evidence that stars die: (a) Supernova explosions are rare in any one galaxy, but each year astronomers see a few erupt in other galaxies. (b) In our own galaxy, large, expanding shells of hot, low-density gas are left behind by the supernova explosions of stars. (c) In contrast, smaller, cooler nebulae are exhaled by the deaths of lower-mass stars much like the sun.



Although the core of helium ash can't generate nuclear energy, it does grow hotter as it contracts because it is converting gravitational energy into thermal energy (see the previous chapter). The rising temperature heats the unprocessed hydrogen just outside the core, hydrogen that was never before hot enough to fuse. Soon, hydrogen fusion begins in a spherical layer or shell around the exhausted core of the star. Like a grass fire burning outward from an exhausted campfire, the hydrogen-fusion shell creeps outward, leaving helium ash behind and increasing the mass of the helium core.

The flood of energy produced by the hydrogen-fusion shell pushes toward the surface, heating the outer layers of the star and forcing them to expand dramatically (■ Figure 15-2). Stars like the sun become giant stars of 10 to 100 solar radii, and the most massive stars become supergiants some 1000 times larger than the sun. This explains the large diameters and low densities of the giant and supergiant stars. In Chapter 13, you learned about the large sizes and low densities of giant and supergiant stars. Now you understand that these stars were once normal main-sequence stars that expanded when hydrogen shell fusion began.

The expansion of its envelope dramatically changes a star's location in the H–R diagram. Just as contraction heats a star, expansion cools it. As the outer layers of gas expand, energy is absorbed in lifting and expanding the gas. The loss of that energy lowers the temperature of the gas. Consequently, the point that represents the star in the H–R diagram moves to the right relatively quickly (in less than a million years for a star of 5 solar masses). A massive star moves to the right across the top of the H–R diagram and becomes a supergiant, while a medium-mass star like the sun becomes a red giant (■ Figure 15-3). As the radius of a giant star continues to increase, its enlarging surface area makes the star more luminous, moving its point upward in

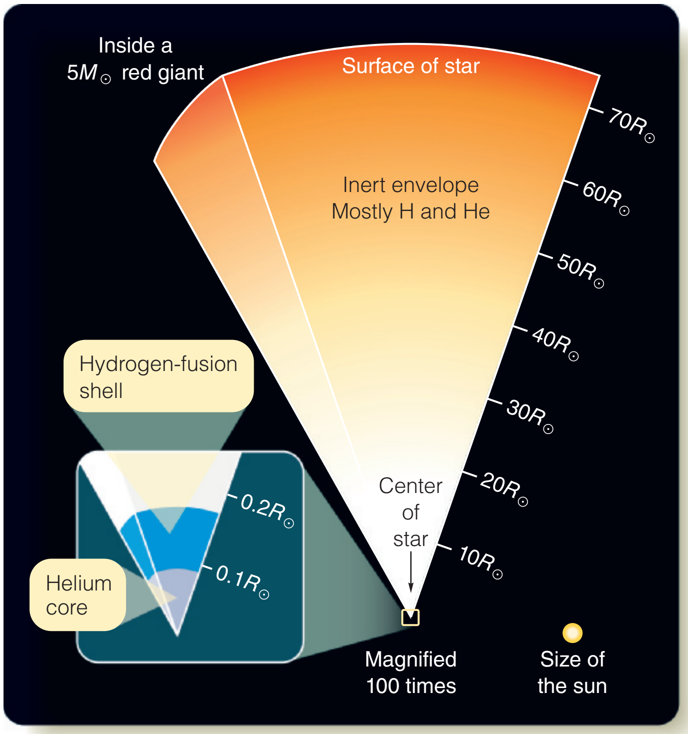


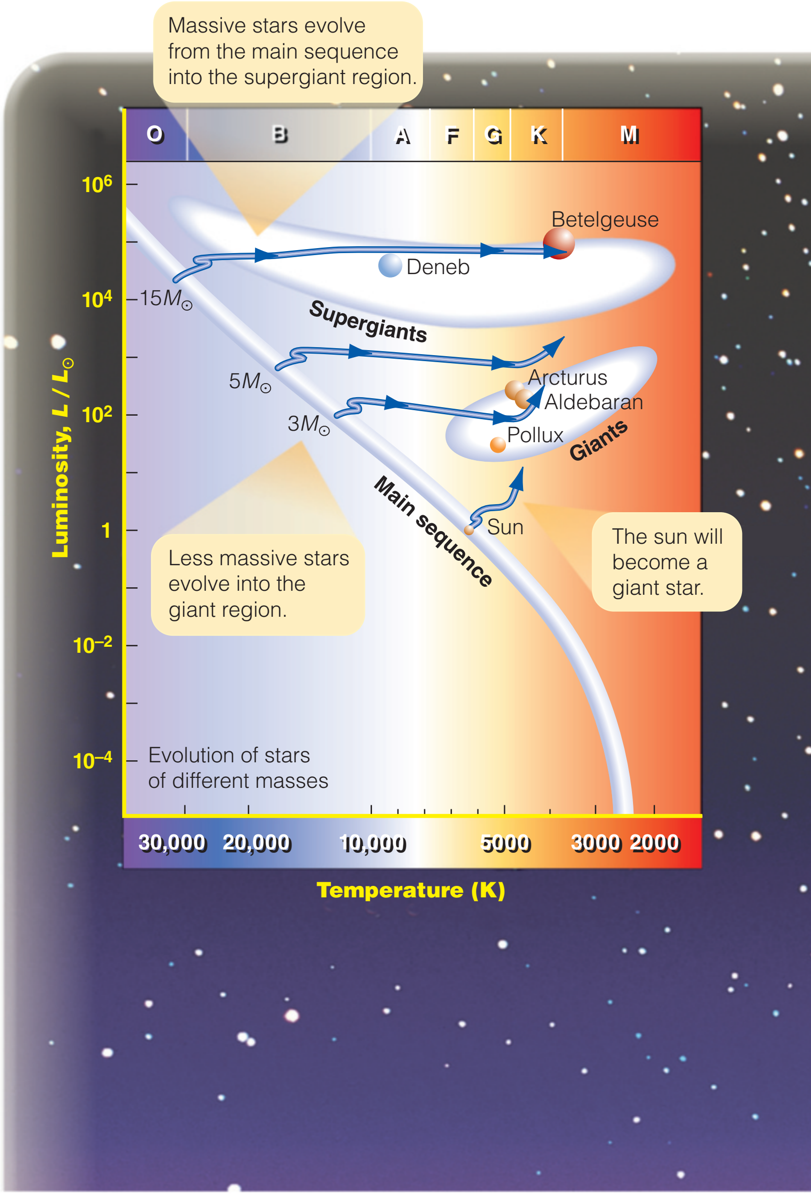
Illustration design by Michael A. Seeds

■ **Figure 15-2**

When a star runs out of hydrogen at its center, the core contracts to a small size, becomes very hot, and begins nuclear fusion in a shell (blue). The outer layers of the star expand and cool. The red giant star shown here has an average density much lower than the air at Earth's surface. Here $M_\odot$ stands for the mass of the sun, and $R_\odot$ stands for the radius of the sun.

■ **Figure 15-3**

The evolution of a massive star moves the point that represents it in the H–R diagram to the right of the main sequence into the region of the supergiants such as Favorite Stars Deneb and Betelgeuse. The evolution of medium-mass stars moves their points in the H–R diagram into the region occupied by giants such as those shown here.



the H–R diagram. Favorite Star Aldebaran, the glowing red eye of Taurus the Bull, is such a red giant, with a diameter 25 times that of the sun but a much cooler surface temperature.

Degenerate Matter

Although the hydrogen-fusion shell can force the envelope of the star to expand, it can't stop the contraction of the helium core. Because the core is not hot enough to fuse helium, gravity squeezes it tighter, and it becomes very small. If you were to represent the helium core of a giant star with a baseball, the outer envelope of the star would be about the size of a baseball stadium. Yet the core would contain about 12 percent of the star's mass. When gas is compressed to such extreme densities, it begins to behave in surprising ways that can affect the evolution of a star. To continue the story of stellar evolution, you need to consider the behavior of gas at extremely high densities.

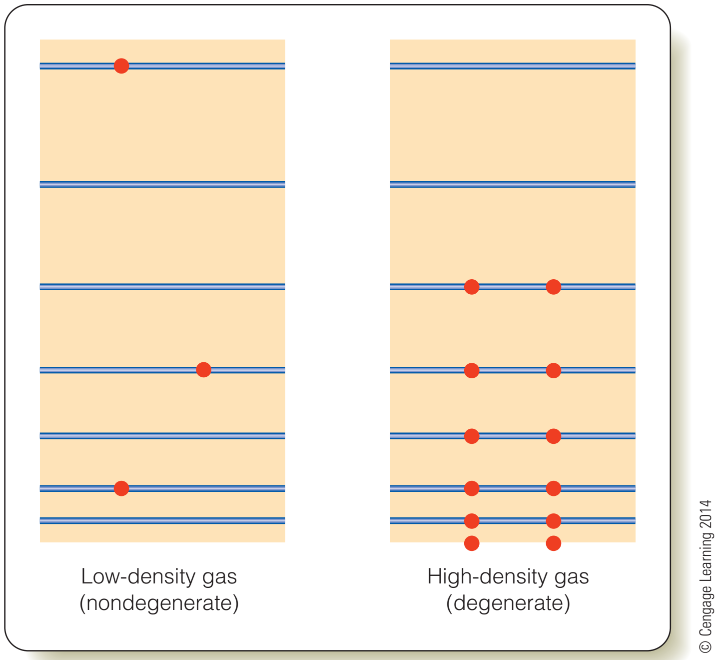
Normally, the pressure in a gas depends on its temperature. The hotter a gas is, the faster its particles move, and the more pressure it exerts. The gas inside a star is ionized, so there are two kinds of particles, atomic nuclei and free electrons. Under normal conditions the gas in a star follows the same laws relating to pressure and temperature as gases do on Earth, but if the gas is compressed to very high densities as in the core of a giant star, two laws of quantum mechanics come into play, and the difference between electrons and nuclei becomes important.

First, quantum mechanics says that the moving electrons confined in the star's core can have only certain amounts of energy, just as the electron in an atom can occupy only certain energy levels (look back to Chapter 6). You can think of these permitted energies as the rungs of a ladder. An electron can occupy any rung but not the spaces between.

The second quantum mechanical law (called the Pauli exclusion principle) says that two identical electrons can't occupy the same energy level. Because electrons spin in one direction or the other, two electrons can occupy a single energy level if they spin in opposite directions. That level is then completely filled, and a third electron can't enter because, whichever way it spins, it will be identical to one or the other of the two electrons already in the level.

A low-density gas has few electrons per cubic centimeter, so there are plenty of energy levels available (■ Figure 15-4). If a gas becomes very dense, however, nearly all of the lower energy levels are occupied. In such a gas, a moving electron can't slow down; slowing down would decrease its energy, and there are no open energy levels for it to drop down to. It can speed up only if it can absorb enough energy to leap to the top of the energy ladder, where there are empty energy levels.

When a gas is so dense that the electrons are not free to change their energy, astronomers call it **degenerate matter.** Although it is a gas, it has two peculiar properties that can affect the star. First, the degenerate gas resists compression. To compress the gas requires pushing against the moving electrons, and changing their motion means changing their energy. That requires tremendous effort because you must boost them to the top of the energy ladder. That is why degenerate matter, though still a gas, is harder to compress than the toughest hardened steel.



■ Figure 15-4

Electron energy levels are arranged like rungs on a ladder. In a low-density gas many levels are open, but in a degenerate gas all lower-energy levels are filled.

Second, the pressure of degenerate gas does not depend on temperature. To see why, note that the pressure depends on the speed of the electrons, which can't be changed without tremendous effort. The temperature, however, depends on the motion of all the particles in the gas, both electrons and nuclei. If you add heat to the gas, most of that energy goes to speed up the motions of the nuclei, which move slowly and don't contribute much to the pressure. Only a few electrons can absorb enough energy to reach the empty energy levels at the top of the energy ladder. That means that changing the temperature of the gas has almost no effect on the pressure.

These two properties of degenerate matter become important when stars end their main-sequence lives (**How Do We Know? 15-1**). Eventually, many stars collapse into white dwarfs, and you will discover that these tiny stars are made of degenerate matter. But long before that, the cores of many giant stars become so dense that they are degenerate, a situation that can produce a cosmic bomb.

Helium Fusion

Hydrogen fusion in main-sequence stars leaves behind helium "ash." Three helium nuclei can collide to make a carbon nucleus in what is called the **triple-alpha process** (nuclear physicists refer to helium nuclei as "alpha particles"). But the helium ash in the core of main-sequence stars is too cool to fuse because helium nuclei

How Do We Know? 15-1

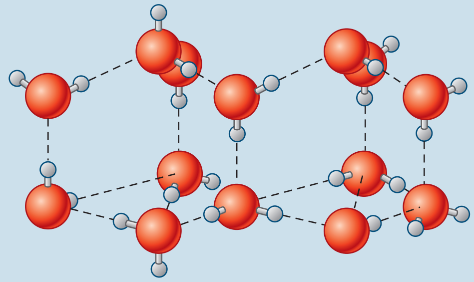
Toward Ultimate Causes

How does a scientist's search for natural causes lead into the world of subatomic particles? Scientists search for causes. They are not satisfied to know that a certain kind of star dies by exploding. They want to know why it explodes. They want to find the causes for the natural events they see, and that search for ultimate causes often leads into the atomic world.

For example, why do icebergs float? When water freezes it becomes less dense than liquid water, so it floats. That answers the question, but you can search for a deeper cause.

Why is ice less dense than water? Water molecules are made up of two hydrogen atoms bonded to an oxygen atom, and the oxygen is so good at attracting electrons, the hydrogen atoms are left needing a bit more negative charge. They are attracted to atoms in nearby molecules. That means the hydrogen atoms in water are constantly trying to stick to other water molecules. When water is warm, the thermal motion prevents these hydrogen bonds from forming, but when water freezes, the hydrogen atoms link the water molecules together. Because of the angles at which the bonds form, the molecules leave open spaces between molecules, and that makes ice less dense than water.

But scientists can continue their search for causes. Why do electrons have negative charge? What is charge? Nuclear particle physicists are tying to understand those properties of matter. Sometimes the properties of very large things such as supernovae are determined by the properties of the tiniest particles. Science is exciting because the simple observation that ice floats in your lemonade can lead you toward ultimate causes and some of the deepest questions about how nature works.



Ice has a low density and floats because of the way electrons (blue) link to oxygen (red) when water freezes.

have positive charges twice that of hydrogen nuclei, so the helium nuclei need to move very fast to overcome that higher Coulomb barrier. As a giant star fuses hydrogen to helium in an expanding shell, its inert core of helium contracts and grows hotter. When the temperature of the core finally reaches 100,000,000 K, it begins to fuse helium nuclei to make carbon.

How a star begins helium fusion depends on its mass. Stars more massive than about 3 solar masses contract rapidly, their helium-rich cores heat up, and helium fusion begins gradually. But less massive stars evolve more slowly, and their cores contract so much that the gas becomes degenerate. On Earth, a teaspoon of the gas would weigh more than an automobile. In this degenerate matter, the pressure does not depend on temperature, and that means the pressure–temperature thermostat does not regulate energy production. When the temperature becomes hot enough, helium fusion begins to make energy and the temperature rises, but pressure does not increase because the gas is degenerate. The higher temperature increases the helium fusion even further, and the result is a runaway explosion called the **helium flash** in which, for a few minutes, the core of a star can generate more energy per second than does an entire galaxy.

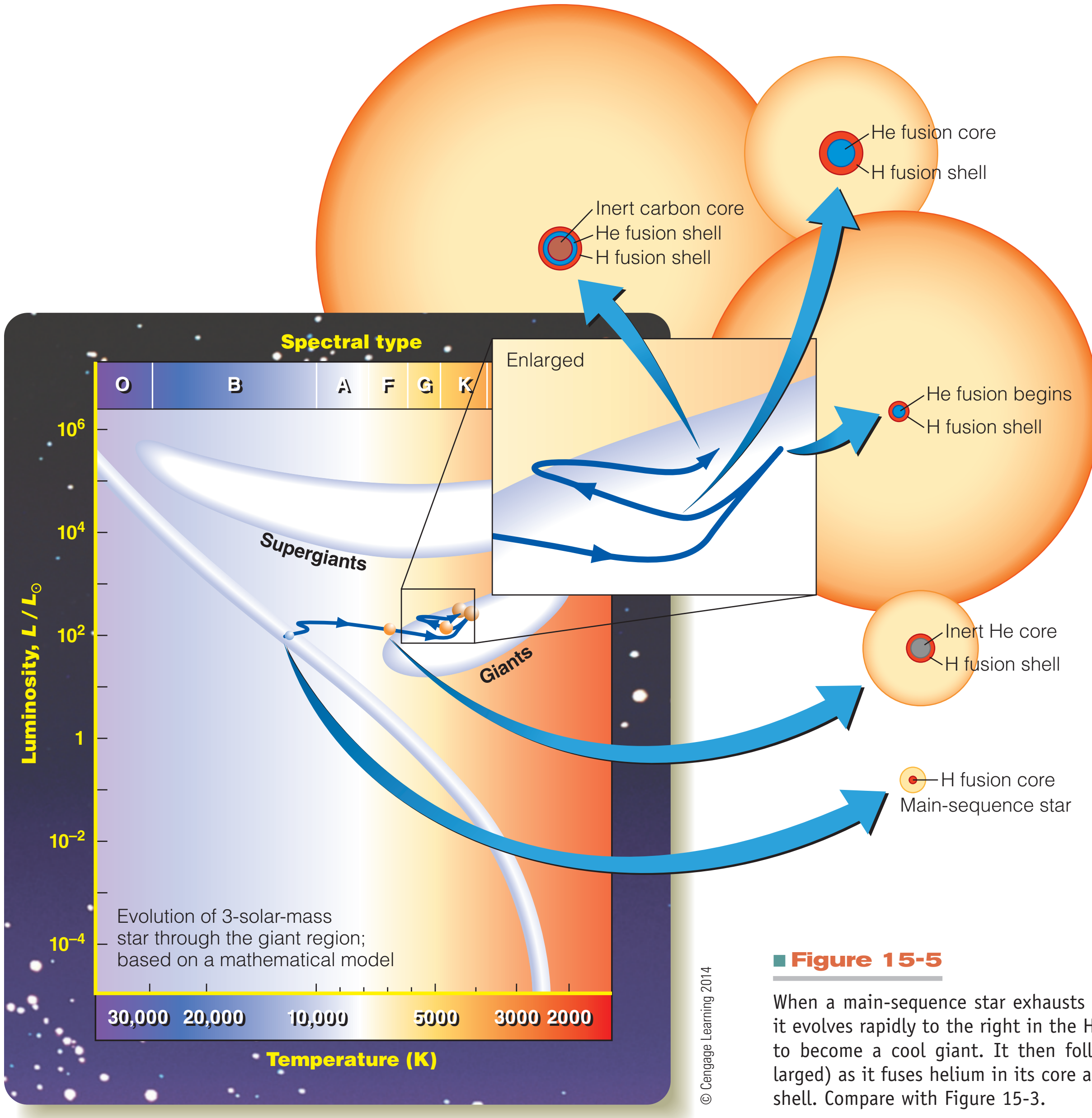
Although the helium flash is sudden and powerful, it does not destroy the star. In fact, if you were observing a giant star as it experienced the helium flash, you would probably see no outward evidence of the eruption. The helium core is quite small (Figure 15-2), and all of the energy of the explosion is absorbed by the distended envelope. In addition, the helium flash is a very short-lived event in the life of a star. In a matter of minutes to hours, the core of the star becomes so hot that a significant number of electrons get boosted to empty energy levels at the top of the energy ladder. That increases the pressure, ends the degenerate conditions, and the pressure–temperature thermostat brings the helium fusion under control. From that point on, the star proceeds to fuse helium steadily in its core.

There are two reasons why you should know about the helium flash. First, it is so violent and so sudden that it makes it difficult to compute models of stars and be sure how they evolve. Astronomers have to exercise ingenuity to get past the helium flash and follow the further evolution of stars. Second, the helium flash is a good illustration of how science reveals a hidden universe. Astronomers would never have known about the helium flash were it not for the theoretical calculation of stellar models.

The sun will experience a helium flash in a few billion years, but stars less massive than about 0.4 solar mass never get hot enough to ignite helium. Stars more massive than 3 solar masses ignite helium before their contracting cores become degenerate.

Whether a star experiences a helium flash or not, the ignition of helium in the core changes the structure of the star. The star now makes energy both in its helium-fusion core and in its hydrogen fusion shell. The energy flowing outward from the core can halt the contraction of the core, and the distended envelope of the star contracts and grows hotter. Consequently, the point that represents the star in the H–R diagram moves downward and to the left toward the hot side of the H–R diagram (■ Figure 15-5).

Helium fusion produces carbon, and some of the carbon nuclei absorb helium nuclei to form oxygen. A few of the oxygen



■ **Figure 15-5**

When a main-sequence star exhausts the hydrogen in its core, it evolves rapidly to the right in the H–R diagram as it expands to become a cool giant. It then follows a looping path (enlarged) as it fuses helium in its core and then fuses helium in a shell. Compare with Figure 15-3.

nuclei can absorb helium nuclei and form neon and then magnesium. Some of these reactions release neutrons, which, having no charge, are more easily absorbed by nuclei to gradually build even heavier nuclei. These reactions are not important as energy producers, but they are slow-cooker processes that form small traces of heavier elements right up to bismuth with atomic weight 209, nearly four times heavier than iron. Many of the atoms in your body were produced this way.

As the helium fuel is used up, the accumulation of carbon and oxygen atoms creates an inert core too cool to fuse. Once again, the core contracts and heats up, and soon a helium-fusion shell ignites below the hydrogen-fusion shell. Now that the star makes energy in two fusion shells, it quickly expands, and its surface cools once again. The point that represents the star in the H–R diagram moves back to the right, completing a loop (Figure 15-5).

What happens to a star after helium fusion depends on its mass, but no matter what tricks the star plays to delay its end, it can't survive long. It must eventually collapse and end its career as a star. The remainder of this chapter will trace the details of this process of stellar death, but before you begin that story, you have to ask the most important question in science: What is the evidence for what you have learned so far? What evidence shows that stars actually evolve as theories predict? You will find the answers in clusters of stars.

Star Clusters: Evidence of Evolution

Just as Sherlock Holmes studies peculiar dust on a lampshade as evidence that will solve a mystery, astronomers look at star clusters and say, "Aha!" A photo of a star cluster freezes a moment in the evolution of the cluster and makes the evolution of the stars visible to human observers.

Because they formed nearly simultaneously from the same gas cloud, the stars in a cluster have about the same age and composition; so any differences you see among them are due to their differences in mass. That means that when you look at a cluster, you can see the effects of stellar evolution as it acts on otherwise similar stars of different mass. Study **Star Cluster H–R Diagrams** on pages 324–325 and notice three important points:

1. There are two kinds of star clusters, *open clusters* and *globular clusters.* They look different, but they are similar in the way their stars evolve. You will learn even more about these clusters in a later chapter.
2. You can estimate the age of a star cluster by observing the *turnoff point* in the distribution of the points that represent its stars in the H–R diagram.
3. Finally, the shape of a star cluster's H–R diagram is governed by the evolutionary path the stars take. The H–R diagrams of older clusters are especially clear in outlining how stars evolve away from the main sequence to the giant region, then move left along the *horizontal branch* before evolving back into the giant region. By comparing clusters of different ages, you can visualize how stars evolve almost as if you were watching a film of a star cluster evolving over billions of years.

If it were not for star clusters, astronomers would have little confidence in the theories of stellar evolution. Star clusters make evolution visible and assure astronomers that they really do understand how stars are born, live, and die.

SCIENTIFIC ARGUMENT

Why is it only lower-mass stars that outline the horizontal branch?

This argument depends on timing. If a star cluster is young, it may contain a few massive stars, but because massive stars are so rare and evolve so rapidly you are unlikely to see more than a few of these stars evolving through the giant or supergiant regions of the H–R diagram. Lower-mass stars are very common and evolve slowly, so in an older star cluster you can see lots of stars in various stages of the post-main-sequence evolution. That outlines the horizontal branch.

Now construct a different argument. **What evidence can you cite that giant stars are main-sequence stars that have expanded to large diameters?**

15-2 The Deaths of Lower-Main-Sequence Stars

Contracting stars heat up by converting gravitational energy into thermal energy. Low-mass stars have little gravitational energy, so when they contract, they don't get very hot. This limits the fuels they can ignite. In the previous chapter, you saw that protostars less massive than 0.08 solar mass can't get hot enough to ignite hydrogen. This section will concentrate on stars more massive than 0.08 solar mass but no more than a few times the mass of the sun.

Structural differences divide the lower-main-sequence stars into two subgroups—very-low-mass red dwarfs and medium-mass stars such as the sun. The critical difference between the two groups is the extent of interior convection. If the star is convective, fuel is constantly mixed, and its resulting evolution is drastically altered.

Red Dwarfs

Stars between 0.08 and about 0.4 solar mass—the red dwarfs—have two advantages over more massive stars. First, they have very small masses, and that means they have very little weight to support. Their pressure–temperature thermostats are set low, and they consume their hydrogen fuel very slowly. The discussion of the life expectancies of stars in the previous chapter concluded that the red dwarfs should live very long lives.

The red dwarfs have a second advantage in that they are totally convective. That is, they are stirred by circulating currents of hot gas rising from the interior and cool gas sinking inward. This means the stars are mixed like a pot of soup that is constantly stirred as it cooks. Hydrogen is consumed uniformly throughout the star, which means the star is not limited to the fuel in its core. It can use all of its hydrogen to prolong its life on the main sequence.

Because a red dwarf is mixed by convection, it can't develop an inert helium core surrounded by unprocessed hydrogen. Then it can never ignite a hydrogen shell and can't become a giant star. What astronomers know about stellar evolution indicates that these red dwarfs should use up nearly all of their hydrogen and live very long lives on the lower main sequence, surviving for a hundred billion years or more. Of course, astronomers can't test this part of their theories because the universe is only 13.7 billion years old, so not a single red dwarf has died of old age anywhere in the universe. Every red dwarf that has ever been born is still shining today.

Medium-Mass Stars

Stars like the sun eventually become hot enough to ignite helium as they pass through their giant phase, but, if they contain less than 4 solar masses,* they do not get hot enough to ignite carbon,

*This mass limit is uncertain, as are many of the masses quoted here. The evolution of stars is highly complex, and such parameters are difficult to specify.

Star Cluster H–R Diagrams

1 An **open cluster** is a collection of 10 to 1000 stars in a region about 25 pc in diameter. Some open clusters are quite small and some are large, but they all have an open, transparent appearance because the stars are not crowded together.

In a star cluster each star follows its orbit around the center of mass of the cluster.

NOAO/AURA/NSF

Visual-wavelength image

Open Cluster
The Jewel Box

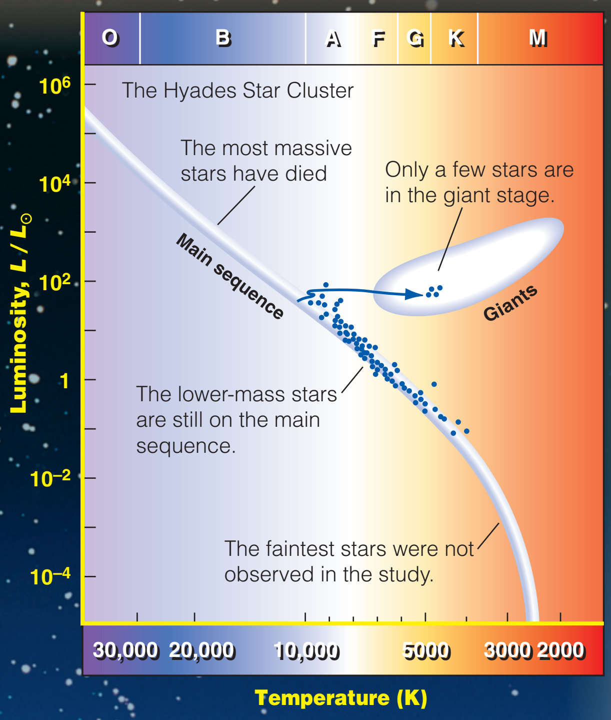
1a A **globular cluster** can contain 10^5 to 10^6 stars in a region only 10 to 30 pc in diameter. The term "globular cluster" comes from the word "globe," although globular cluster is pronounced like "glob of butter." These clusters are nearly spherical, and the stars are much closer together than the stars in an open cluster.

Globular Cluster
47 Tucanae

Australian Astronomical Observatory/David Malin

Visual-wavelength image

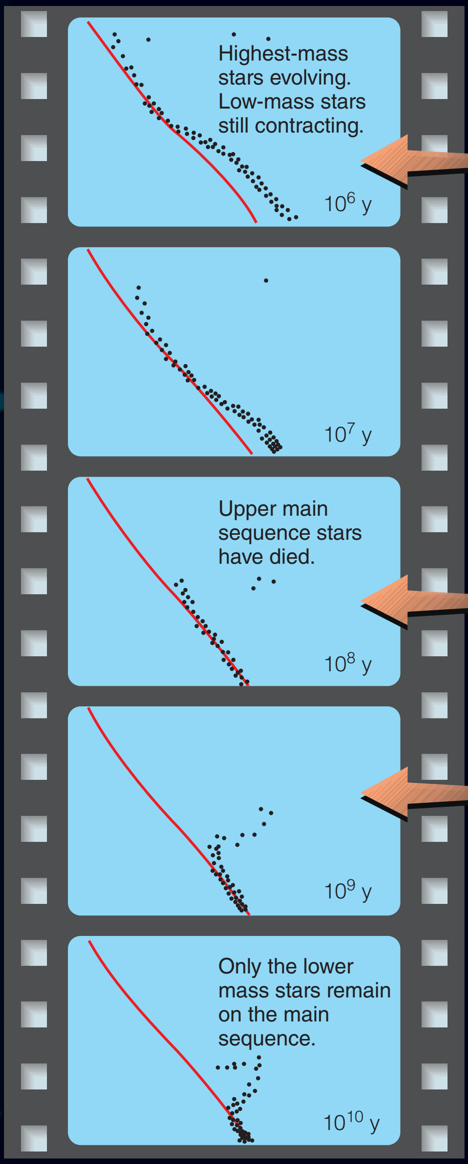
Astronomers can construct an H–R diagram for a star cluster by plotting a point to represent the luminosity and temperature of each star.

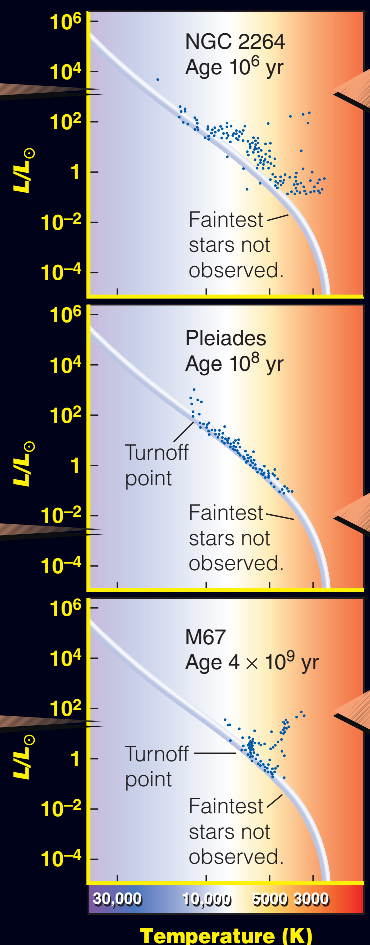


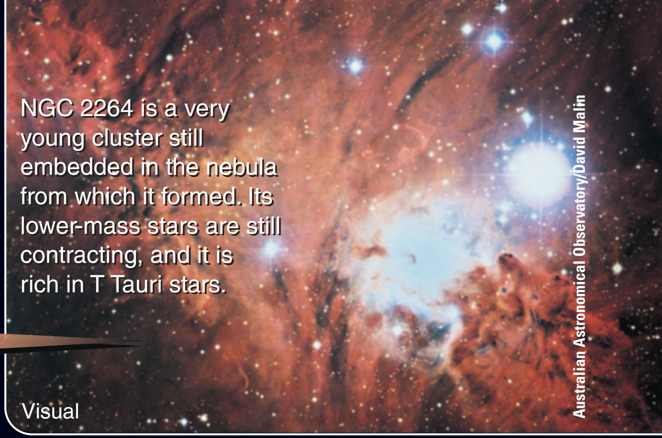
2 The H–R diagram of a star cluster can make the evolution of stars visible. The key is to remember that all of the stars in the star cluster have the same age but differ in mass. The H–R diagram of a star cluster provides a snapshot of the evolutionary state of the stars at the time you happen to be alive. The diagram here shows the 650-million-year-old star cluster called the Hyades. The upper main sequence is missing because the more massive stars have died, and our snapshot catches a few medium-mass stars leaving the main sequence to become giants.

As a star cluster ages, its main sequence grows shorter like a candle burning down. You can judge the age of a star cluster by looking at the **turnoff point**, the point on the main sequence where stars evolve to the right to become giants. Stars at the turnoff point have lived out their lives and are about to die. Consequently, the life expectancy of the stars at the turnoff point equals the age of the cluster.

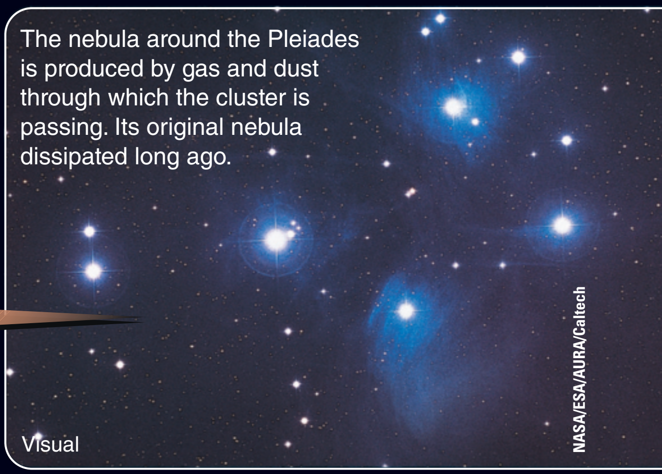
3 From theoretical models of stars, you could construct a film to show how the H–R diagram of a star cluster changes as it ages. You can then compare theory (left) with observation (right) to understand how stars evolve. Note that the time step for each frame in this film increases by a factor of 10.



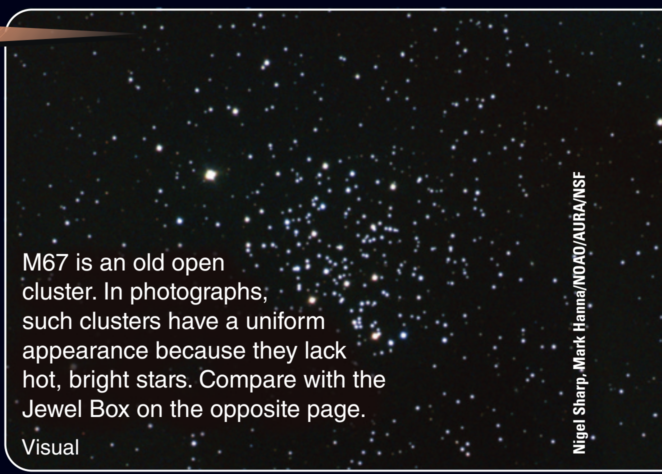




NGC 2264 is a very young cluster still embedded in the nebula from which it formed. Its lower-mass stars are still contracting, and it is rich in T Tauri stars.

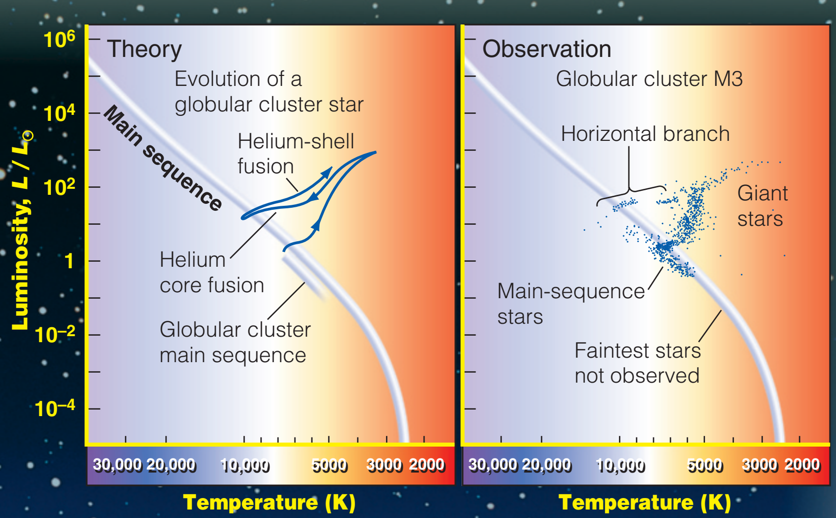


The nebula around the Pleiades is produced by gas and dust through which the cluster is passing. Its original nebula dissipated long ago.



M67 is an old open cluster. In photographs, such clusters have a uniform appearance because they lack hot, bright stars. Compare with the Jewel Box on the opposite page.

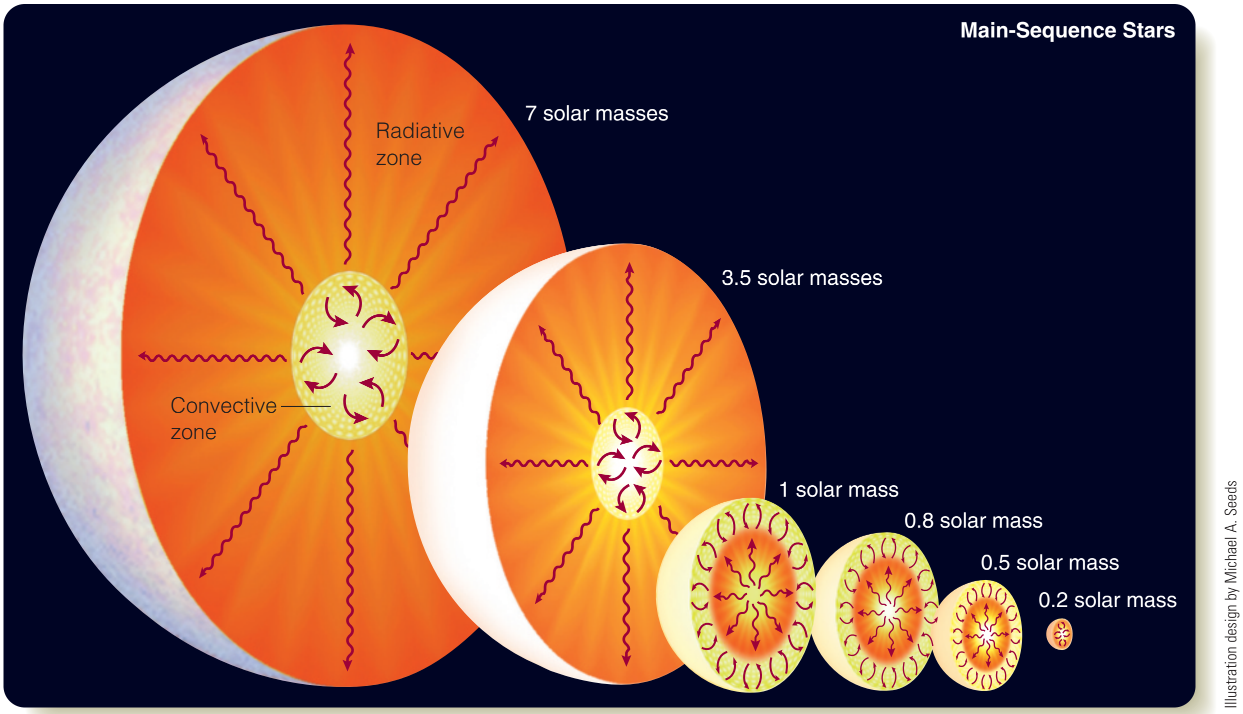
Globular cluster H–R diagrams resemble the last frame in the film, which tells you that globular clusters are very old.



3a The H–R diagrams of globular clusters have very faint turnoff points showing that they are very old clusters. The best analysis suggests these clusters are about 11 billion years old.

The **horizontal branch** stars are giants fusing helium in their cores and then in shells. The shape of the horizontal branch outlines the evolution of these stars.

The main-sequence stars in globular clusters are fainter and bluer than the zero-age main sequence. Spectra reveal that globular cluster stars are poor in elements heavier than helium, and that means their gases are less opaque. That means energy can flow outward more easily, which makes the stars slightly smaller and hotter. Again the shape of star cluster H–R diagrams illustrates principles of stellar evolution.



■ **Figure 15-6**

Inside main-sequence stars. The more massive stars have small convective interiors and radiative envelopes. Stars like the sun have radiative interiors and convective envelopes. The lowest-mass stars are convective throughout. The "cores" of the stars where nuclear fusion occurs (not shown) are smaller than the interiors.

the next fuel after helium. When they reach that impasse, they collapse and become white dwarfs. There are two keys to the evolution of these sunlike stars, the lack of complete mixing and mass loss.

The interiors of medium-mass stars are not completely mixed (■ Figure 15-6). Stars of 1.1 solar masses or less have no convection near their centers, so they are not mixed at all. Stars with a mass greater than 1.1 solar masses have small zones of convection at their centers, but this mixes no more than about 12 percent of the star's mass. Medium-mass stars, whether they have convective cores or not, are not thoroughly mixed, and the helium ash accumulates in an inert helium core surrounded by unprocessed hydrogen. Recall from earlier in this chapter that when this core contracts the unprocessed hydrogen ignites in a shell and swells the star into a giant.

In the giant stage, the core of the star contracts, and the envelope expands. The star fuses helium first in its core and then in an expanding shell surrounding a core of carbon and oxygen. This core contracts and grows hotter, but because the star has too low a mass the core can't get hot enough to ignite carbon fusion. The carbon-oxygen core is a dead end for these medium-mass stars.

All of this discussion is based on theoretical models of stars and a general understanding of how stars evolve. Does it really happen? Astronomers need observational evidence to confirm their theories, and the gas that is expelled from these giant stars gives visible evidence that sunlike stars do indeed die in this way.

Planetary Nebulae

When a medium-mass star like the sun becomes a distended giant, its atmosphere cools. As it cools, it becomes more opaque, and light has to push against it to escape. At the same time the fusion shells become so thin they are unstable and begin to flare, which also pushes the atmosphere outward. Because of this outward pressure, an aging giant can expel its outer atmosphere in repeated surges to form one of the most interesting objects in astronomy, a **planetary nebula,** so called because through a small telescope it looks like the greenish-blue disk of a planet like Uranus or Neptune. In fact, a planetary nebula has nothing to do with a planet. It is composed of ionized gases expelled by a dying star.

Study **The Formation of Planetary Nebulae** on pages 328–329 and notice four things:

1. You can understand what planetary nebulae are like by using simple observational principals such as Kirchhoff's laws and the Doppler effect.

2 Notice the model that astronomers have developed to explain planetary nebulae. The real nebulae are more complex than the simple model of a slow wind and a fast wind, but the model provides a way to organize the observed phenomena.

3 Oppositely directed jets (much like bipolar flows from protostars) produce many of the asymmetries seen in planetary nebulae.

4 The star itself must contract into a white dwarf.

Most astronomy books say that the sun will form a planetary nebula, but that may not happen. To ionize the gas and light up a planetary nebula, a star must become a white dwarf with a temperature of at least 25,000 K. Mathematical models show that a collapsing star of less than 0.55 solar mass can take as long as a million years to heat up enough to ionize its nebulae, and by that time the expelled gases are long gone. Models of the sun are not precise enough to indicate how much mass will be left once it ejects its outer layers. If it is left with too little mass, it may heat too slowly. Also, some research suggests that a star needs a binary companion to speed up its spin and make it eject a planetary nebula. The sun, of course, has no binary companion. This is an area of active research, and there are no firm conclusions. Are you disappointed that the sun may not light up its own planetary nebula? At least that potential embarrassment lies a few billion years in the future.

Medium-mass stars die by ejecting gas into space and contracting into white dwarfs. You have found evidence regarding the deaths of medium-mass stars in observations of planetary nebulae. Now you can turn your attention to the evidence revealed by white dwarfs themselves.

White Dwarfs

When you surveyed the stars (look back to Chapter 13), you discovered that white dwarfs are the second most common kind of star. Only red dwarfs are more abundant. Now you can recognize the billions of white dwarfs in our galaxy as the remains of medium-mass stars.

The first white dwarf discovered was the faint companion to Favorite Star Sirius. In that visual binary system, the bright star is Sirius A. The white dwarf, Sirius B, is 10,000 times fainter than Sirius A. The orbital motions of the stars (shown in Figure 13-14) reveal that the white dwarf contains 0.98 solar mass, and its blue-white color tells you that its surface is hot, about 45,000 K. Because it is both very hot and very low luminosity, it must have a small surface area (see Reasoning with Numbers 13-3)—in fact, it is about the size of Earth. Dividing its mass by its volume reveals that it is very dense—about 2×10^6 g/cm^3. On Earth, a teaspoonful of Sirius B material would weigh more than 11 tons.

A normal star is supported by energy flowing outward from its core, but a white dwarf can't generate energy by nuclear fusion. It has exhausted its hydrogen and helium fuels and converted them into carbon and oxygen. When a star collapses into a white dwarf, it converts gravitational energy into thermal energy. Its interior becomes very hot, but it can't get hot enough to fuse its carbon-oxygen interior. Instead the star contracts until it becomes degenerate. Although a tremendous amount of energy flows out of the hot interior, it is not the energy flow that supports the star. The white dwarf is supported against its own gravity by the pressure of its degenerate electrons.

The interior of a white dwarf is mostly carbon and oxygen nuclei immersed in a whirling storm of degenerate electrons. Theory predicts that as the star cools these particles will lock together to form a crystal lattice, so there may be some truth in thinking of the interiors of aging white dwarfs as great crystals of carbon and oxygen. Near the surface, where the pressure is lower, a layer of ionized gases makes up a hot atmosphere. A 150-lb human would weigh 50 million pounds on the surface of a white dwarf. That strong gravity pulls the atmosphere down into a shallow layer. If Earth's atmosphere were equally shallow, people on the top floors of skyscrapers would have to wear space suits.

Clearly, a white dwarf is not a true star. It generates no nuclear energy, is almost totally degenerate, and, except for a thin layer at its surface, contains no gas. Instead of calling a white dwarf a "star," you can call it a "compact object." The next chapter discusses two other compact objects, neutron stars and black holes.

A white dwarf's future is bleak. As it radiates energy into space, its temperature gradually falls, but it can't shrink any smaller because its degenerate electrons resist getting closer together. Degenerate matter is a very good thermal conductor, so as heat flows to the surface and escapes into space the white dwarf gets fainter and cooler, moving downward and to the right in the H–R diagram. Because a white dwarf contains a tremendous amount of heat, it needs billions of years to radiate that heat through its small surface area. Eventually, such objects may become cold and dark, so-called **black dwarfs.** Our galaxy is not old enough to contain black dwarfs. The coolest white dwarfs in our galaxy are about the temperature of the sun.

Perhaps the most interesting thing astronomers have learned about white dwarfs has come from mathematical models. The equations predict that if you add mass to a white dwarf, its radius will *shrink* because added mass will increase its gravity and squeeze it tighter. If you add enough to raise its total mass to about 1.4 solar masses, the equations predict that its radius will shrink to zero (■ Figure 15-7). This is called the **Chandrasekhar limit** after Subrahmanyan Chandrasekhar, the astronomer who discovered it (Subrahmanyan was his family name, Chandrasekhar his given name). That limit seems to imply that a star more massive than 1.4 solar masses can't become a white dwarf unless it sheds mass in some way.

The Formation of Planetary Nebulae

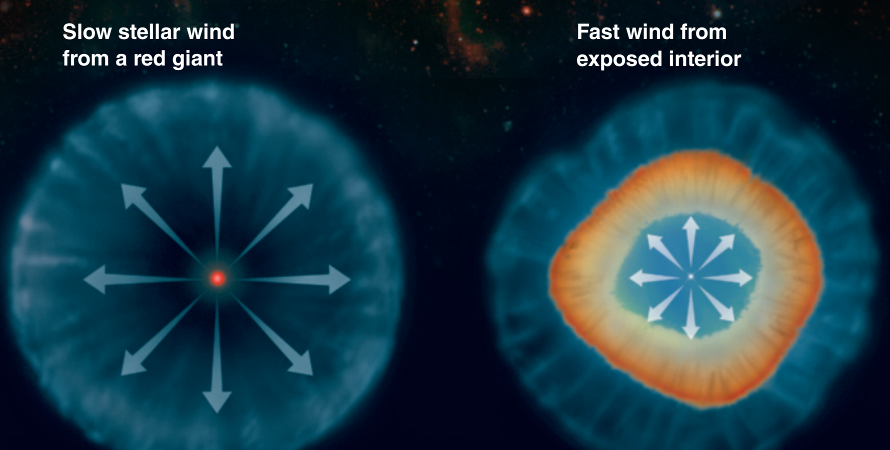
1 Simple observations tell astronomers what planetary nebulae are like. Their angular size and their distances indicate that their radii range from 0.2 to 3 ly. The presence of emission lines in their spectra assures that they are excited, low-density gas. Doppler shifts show they are expanding at 10 to 20 km/s. If you divide radius by velocity, you find that planetary nebulae are no more than about 10,000 years old. Older nebulae evidently become mixed into the interstellar medium.

Astronomers find about 1500 planetary nebulae in the sky. Because planetary nebulae are short-lived formations, you can conclude that they must be a common part of stellar evolution. Medium-mass stars up to a mass of about 8 solar masses are destined to die by forming planetary nebulae.

NASA/JPL-Caltech/ESA

The Helix Nebula is 2.5 ly in diameter, and the radial texture shows how light and winds from the central star are pushing outward.

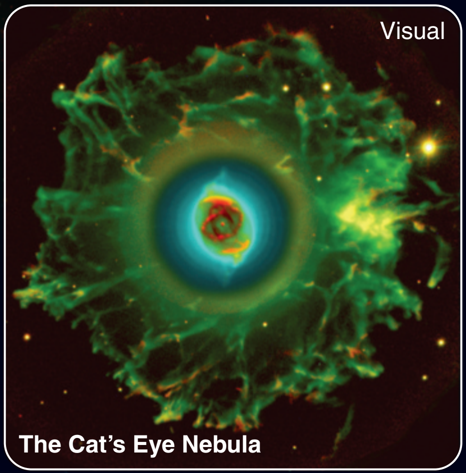
2 The process that produces planetary nebulae involves two stellar winds. First, as an aging giant, the star gradually blows away its outer layers in a slow breeze of low-excitation gas that is not easily visible. Once the hot interior of the star is exposed, it ejects a high-speed wind that overtakes and compresses the gas of the slow wind like a snowplow, while ultraviolet radiation from the hot remains of the central star excites the gases to glow like a giant neon sign.



The gases of the slow wind are not easily detectable.

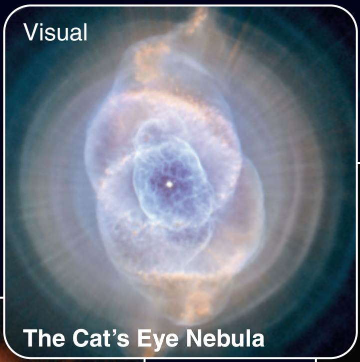
You see a planetary nebula where the fast wind compresses the slow wind.

2a The Cat's Eye, below, lies at the center of an extended nebula that must have been exhaled from the star long before the fast wind began forming the visible planetary nebula. See other images of the nebula on opposite page.



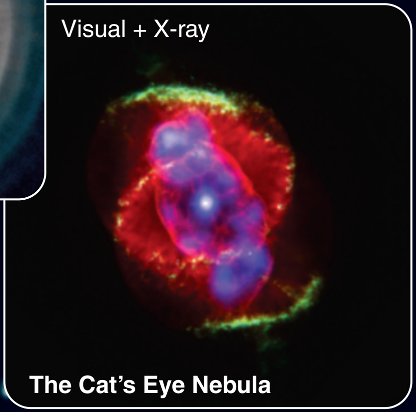
Roman Corradi/Nordic Optical Telescope

3 Images from the Hubble Space Telescope reveal that asymmetry is the rule in planetary nebulae rather than the exception. A number of causes have been suggested. A disk of gas around a star's equator might form during the slow-wind stage and then deflect the fast wind into oppositely directed flows. Another star or planets orbiting the dying star, rapid rotation, or magnetic fields might cause these peculiar shapes. The Hour Glass Nebula seems to have formed when a fast wind overtook an equatorial disk (white in the image). The nebula Menzel 3, as do many planetary nebulae, shows evidence of multiple ejections.


Visual
The Cat's Eye Nebula

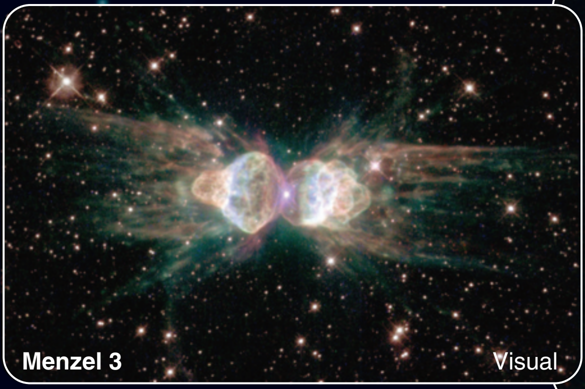
NASA, ESA, HEIC, and The Hubble Heritage Team (STScI/AURA)

Some shapes suggest bubbles being inflated in the interstellar medium. The Cat's Eye is shown at left, below, and on the facing page.

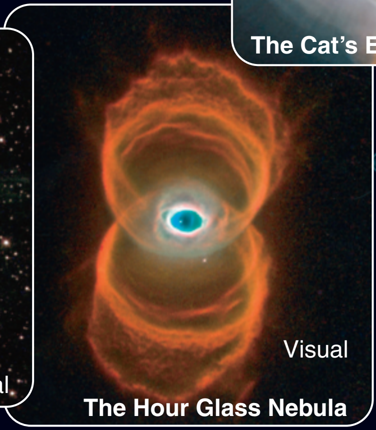

Visual + X-ray
The Cat's Eye Nebula

NASA/UIUC/Y.Chu et al.

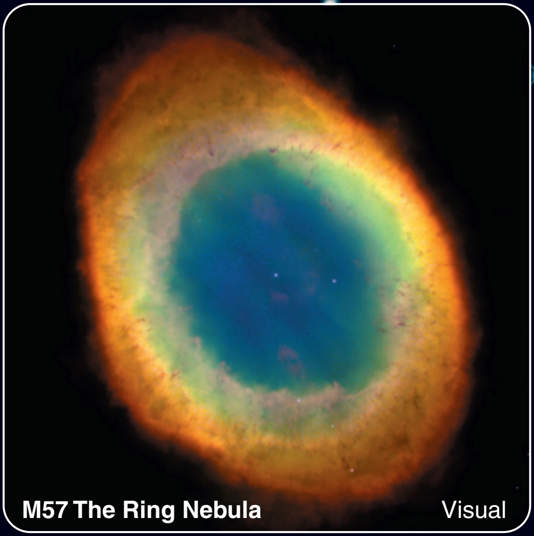
The purple glow in the image above is a region of X-ray bright gas with a temperature measured in millions of degrees. It is apparently driving the expansion of the nebula.


Menzel 3 Visual

STScI AVL


Visual
The Hour Glass Nebula

R. Sahai and J. Trauger (JPL), WFPC2 Science Team, NASA

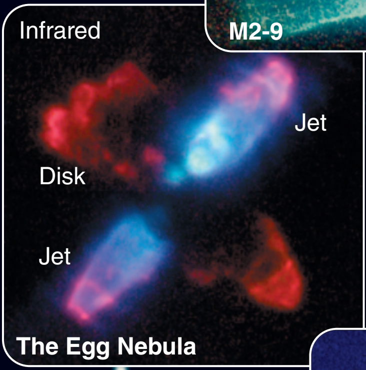

M57 The Ring Nebula Visual

The Hubble Heritage Team (AURA/STScI/NASA)

Some planetary nebulae, such as M2-9, at right, are highly elongated, and it has been suggested that the Ring Nebula, at left, is a tubular shape that happens to be pointed roughly at Earth.

NASA, Bruce Balick University of Washingtom, Vincent Icke Leiden

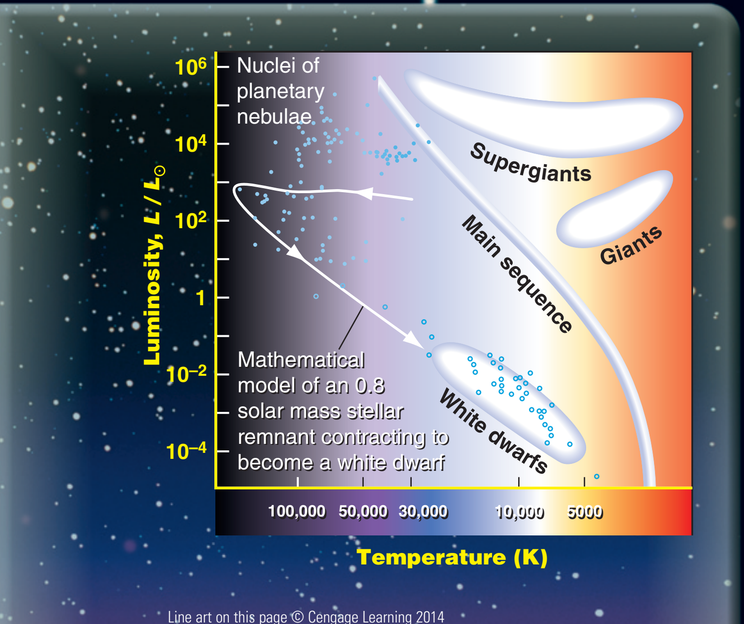

M2-9 Visual



Infrared
The Egg Nebula

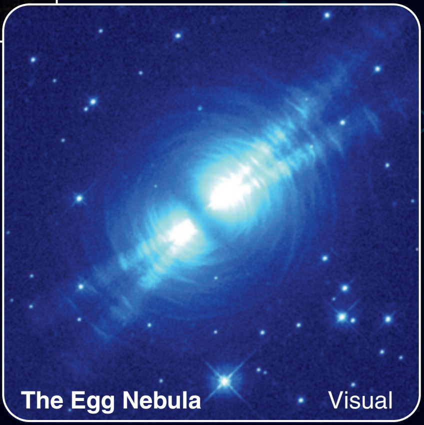
Rodger Thompson, Marcia Rieke, Glenn Schneider, Dean Hines (University of Arizona); Raghvendra Sahai (Jet Propulsion Laboratory); NICMOS Instrument Definition Team, and NASA

At visual wavelengths, the Egg Nebula is highly elongated, as shown below. The infrared image at left reveals an irregular, thick disk from which jets of gas and dust emerge. Such beams may create many of the asymmetries in planetary nebulae.



Line art on this page © Cengage Learning 2014

4 Once an aging giant star blows its surface into space to form a planetary nebula, the remaining hot interior collapses into a small, intensely hot object containing a carbon and oxygen interior surrounded by hydrogen and helium fusion shells and a thin atmosphere of hydrogen. The fusion gradually dies out, and the core of the star evolves to the left of the conventional H–R diagram to become the intensely hot nucleus of a planetary nebulae. Mathematical models show that these nuclei cool slowly to become white dwarfs.


The Egg Nebula Visual

Rodger Thompson, Marcia Rieke, Glenn Schneider, Dean Hines (University of Arizona); Raghvendra Sahai (Jet Propulsion Laboratory); NICMOS Instrument Definition Team, and NASA

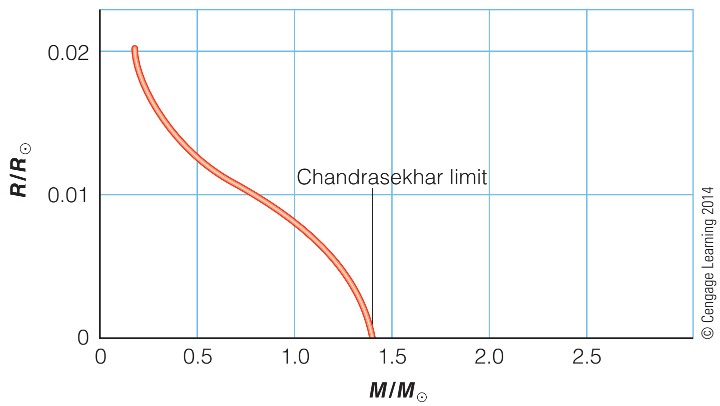


Figure 15-7

The more massive a white dwarf is, the smaller its radius. Stars more massive than the Chandrasekhar limit of 1.4 $M_{\odot}$ can't be white dwarfs.

Stars do lose mass. Observations provide clear evidence that young stars have strong stellar winds, and aging giants and supergiants also lose mass rapidly (■ Figure 15-8). This suggests that stars more massive than the Chandrasekhar limit can eventually die as white dwarfs if they reduce their mass. Theoretical models show that stars that begin life with as much as 8 solar masses should lose mass fast enough to reduce their mass below 1.4 solar masses and eventually collapse to form white dwarfs. With mass loss, a wide range of medium-mass stars can eventually die as white dwarfs.

SCIENTIFIC ARGUMENT

What evidence can you cite to show that large numbers of stars die by producing planetary nebula?

You can begin your argument by noting that planetary nebulae are only a light-year or so in radius and that Doppler shifts show that they are expanding at 10 to 20 km/s. Dividing the radius by the velocity, tells you that a typical planetary nebula is only about 10,000 years old. That means that the nebulae don't last very long. Nevertheless, astronomers find 1500 of them visible in the sky. To be so common but so short-lived, planetary nebulae must be produced in large numbers as medium-mass stars blow their outer layers into space.

Now review more evidence. **Use Favorite Star Sirius to explain how you know that white dwarfs are very dense.**

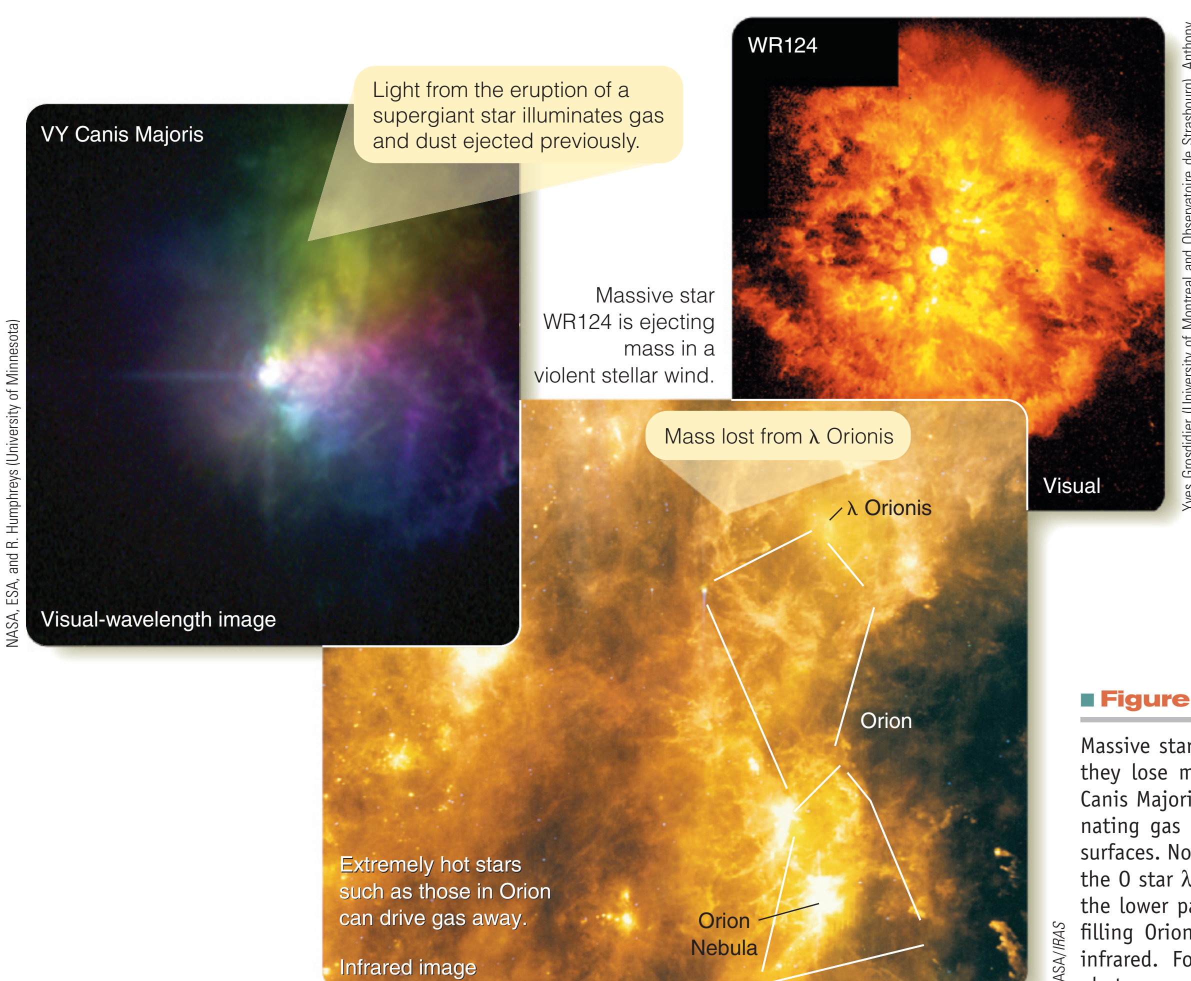


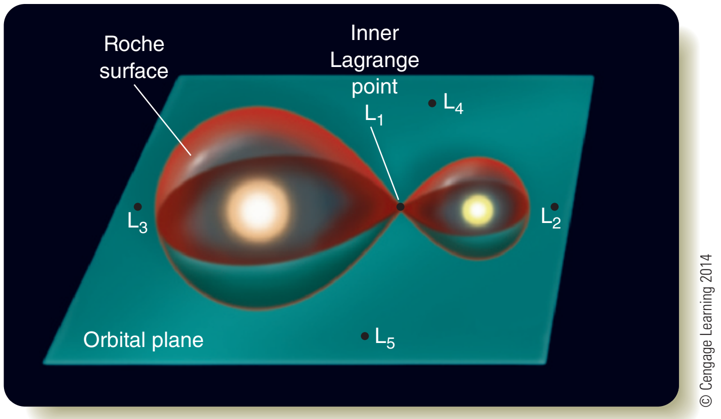
Figure 15-8

Massive stars can be so hot and so large they lose mass. The two supergiants VY Canis Majoris and WR124 are both illuminating gas they have ejected from their surfaces. Note the expanding shell around the O star λ Orionis in the upper right of the lower panel. Warmed dust in the gas filling Orion makes the gas glow in the infrared. For another example, see the photo on page 317.

15-3 The Evolution of Binary Systems

STARS IN BINARY SYSTEMS can evolve independently of each other if their orbits are large. In this situation, one of the stars can swell into a giant and collapse without disturbing its companion. But some binary stars orbit as close to each other as 0.1 AU, and when one of those stars begins to swell into a giant its companion can suffer in peculiar ways.

These interacting binary stars are interesting in their own right. The stars share a complicated history and can experience strange and violent phenomena as they evolve. But such systems are also important because they can help astronomers understand the ultimate fate of stars. In the next chapter you will see how astronomers use interacting binary stars to search for black holes.



■ **Figure 15-9**

A pair of binary stars control the region of space located inside the Roche surface. The Lagrange points are locations of stability, with the inner Lagrange point making a connection through which the two stars can transfer matter.

Mass Transfer

Binary stars can sometimes interact by transferring mass from one star to the other. The gravitational fields of the two stars, combined with the rotation of the binary system, define a dumbbell-shaped volume around the pair of stars called the **Roche lobes.** The surface of this volume is called the **Roche surface.** The size of the Roche lobes depends on the mass of the stars and on the distance between the stars. If the stars are far apart, the lobes are very large, and the stars easily control their own mass. If the stars are close together, however, the lobes are small and can interfere with the evolution of the stars. Matter inside each star's Roche lobe is gravitationally bound to the star, but matter that leaves a star's Roche lobe can fall into the other star or leave the binary system completely.

The **Lagrange points** are places in the orbital plane of a binary star system where a bit of matter can reach stability. For astronomers, the most important of these points is the **inner Lagrange point** where the two Roche lobes meet (■ Figure 15-9). If matter can leave a star and reach the inner Lagrange point, it can flow onto the other star. Thus, the inner Lagrange point is the connection through which the stars can transfer matter.

In general, there are only two ways matter can escape from a star and reach the inner Lagrange point. First, if a star has a strong stellar wind, some of the gas blowing away from it can pass through the inner Lagrange point and be captured by the other star. Second, if an evolving star expands so far that it fills its Roche lobe, which can occur if the stars are close together and the lobes are small, then matter can overflow through the inner Lagrange point onto the other star. Mass transfer driven by a stellar wind tends to be slow, but mass can be transferred rapidly by an expanding star.

Evolution with Mass Transfer

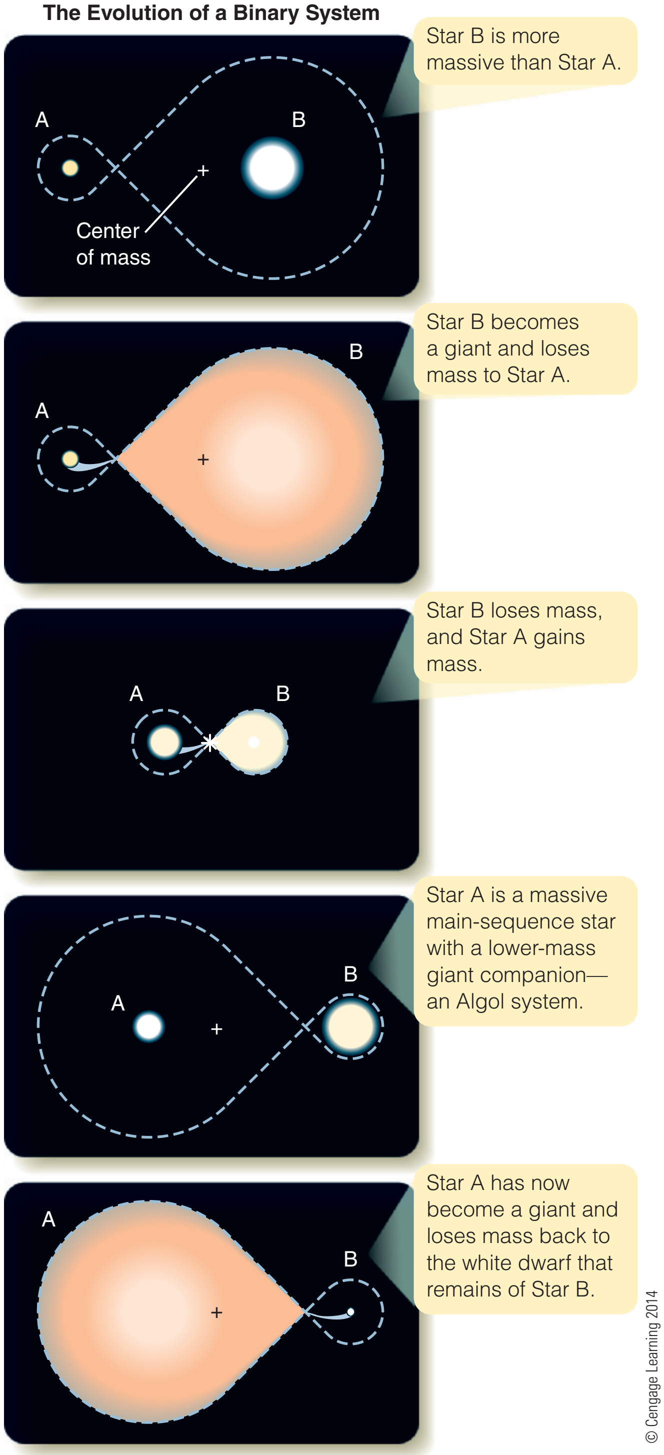
Mass transfer between stars can affect the evolution of the stars in surprising ways. In fact, it is the solution to a problem that puzzled astronomers for many years.

In some binary systems, the less massive star has become a giant, while the more massive star is still on the main sequence. If higher-mass stars evolve faster than lower-mass stars, how do the lower-mass stars in such binaries manage to leave the main sequence first? This is called the Algol paradox, after the binary system Algol (look back to Figure 13-18).

Mass transfer explains how this could happen. Imagine a binary system that contains a 5-solar-mass star and a 1-solar-mass companion. The two stars formed at the same time, so the higher-mass star evolving faster leaves the main sequence first. When it expands into a giant, however, it fills its Roche lobe and transfers matter to the low-mass companion. The higher-mass star loses mass and evolves into a lower-mass star, and the companion gains mass and becomes a higher-mass star that is still on the main sequence. This explains how there could be a system such as Algol that contains a 5-solar-mass main-sequence star and a 1-solar-mass giant.

The first four frames of ■ Figure 15-10 show mass transfer producing a system like Algol. The last frame shows an additional stage in which the giant star has collapsed to form a white dwarf, and the more massive companion has expanded and is transferring matter back to the white dwarf. Such systems can become the site of tremendous explosions. To see how this can happen, you need to think about how mass falls into a star.

The Evolution of a Binary System



■ **Figure 15-10**

A pair of stars orbiting close to each other can exchange mass and modify their evolution.

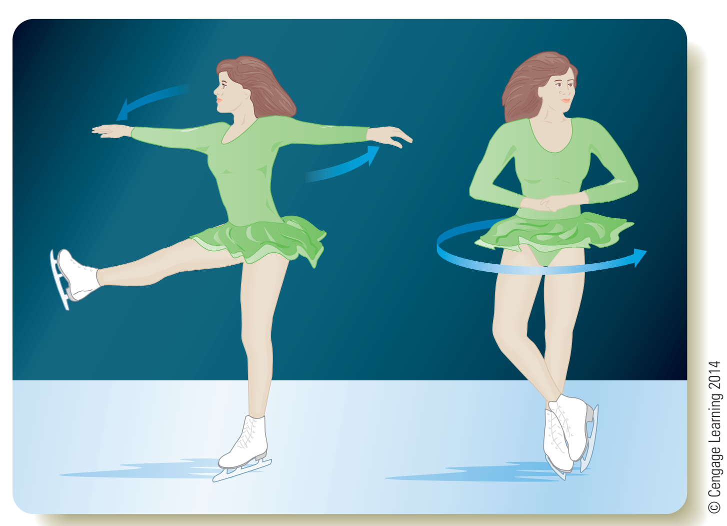
Accretion Disks

Matter flowing from one star to another can't fall directly into the star. Rather, because of conservation of angular momentum, it must flow into a whirling disk around the star.

Angular momentum refers to the tendency of a rotating object to continue rotating. All rotating objects possess some angular momentum, and in the absence of external forces an object maintains (conserves) its total angular momentum. An ice skater takes advantage of conservation of angular momentum by starting a spin slowly with her arms extended and then drawing them in. As her mass becomes concentrated closer to her axis of rotation, she spins faster (■ Figure 15-11). The same effect causes the slowly circulating water in a bathtub to spin in a whirlpool as it approaches the drain.

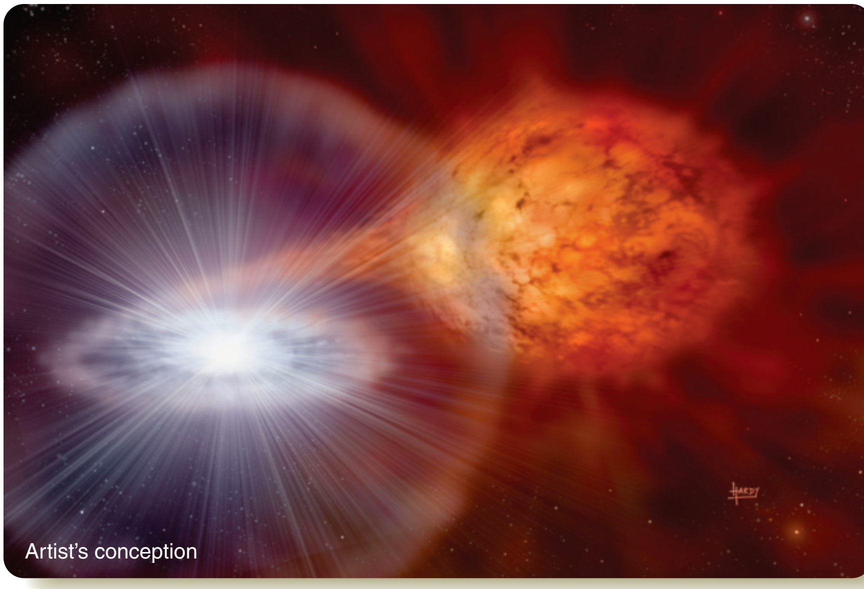
Mass transferred through the inner Lagrange point in a binary system toward a star must conserve its angular momentum. If the star is small enough, as in the case of a white dwarf, the mass will form a rapidly rotating whirlpool called an **accretion disk** (■ Figure 15-12).

Two important things happen in an accretion disk. First, the gas in the disk grows very hot due to friction and tidal forces. The disk also acts as a brake, shifting angular momentum outward in the disk and allowing the innermost matter to fall into the white dwarf. The interior parts of an accretion disk around a white dwarf are violent places. The temperature of the gas can exceed a million Kelvin, causing the gas to emit X-rays, and the matter falling inward can produce a vast explosion when enough accumulates on the white dwarf.



■ **Figure 15-11**

A skater demonstrates conservation of angular momentum when she spins faster by drawing her arms and legs closer to her axis of rotation.



■ **Figure 15-12**

Matter from an evolving red giant falls into a white dwarf and forms a whirling accretion disk. Friction and tidal forces can make the disk very hot. Such systems can lead to nova explosions on the surface of the white dwarf as shown in this artist's impression.

Novae

At the beginning of this chapter you read that the word *nova* refers to what seems to be a new star appearing in the sky for a while and then fading away. Modern astronomers know that a nova is not a new star but an old star flaring up. After a nova fades, astronomers can photograph the spectrum of the remaining faint point of light. Invariably, they find a short-period spectroscopic binary containing a normal star and a white dwarf. A nova is evidently an explosion involving a white dwarf.

Observational evidence can tell you how nova explosions occur. As the explosion begins, spectra show blueshifted absorption lines, which tells you the gas is dense and coming toward you at a few thousand kilometers per second. After a few days, the spectral lines change to emission lines, telling you the gas has thinned. The blueshifts remain, so you can conclude that an expanding cloud of debris has been ejected into space.

Nova explosions occur when mass transfers from a normal star through the inner Lagrange point into an accretion disk around the white dwarf. As the matter loses its angular momentum in the inner accretion disk, it settles onto the surface of the white dwarf and forms a layer of unused nuclear fuel—mostly hydrogen. As the layer deepens, it becomes denser and hotter until the hydrogen fuses in a sudden explosion that blows the surface off the white dwarf. Although the expanding cloud of debris contains less than 0.0001 solar mass, it is hot, and its expanding surface area makes it very luminous. Nova explosions can become 100,000 times more luminous than the sun. As the debris cloud expands, cools, and thins over a period of weeks and months, the nova fades from view.

The explosion of this material hardly disturbs the white dwarf and its companion star. Mass transfer quickly resumes, and a new layer of fuel begins to accumulate. How fast the fuel builds up depends on the rate of mass transfer. You can expect novae to repeat each time an explosive layer accumulates. Many novae take thousands of years to build an explosive layer, but some take only decades (■ Figure 15-13).

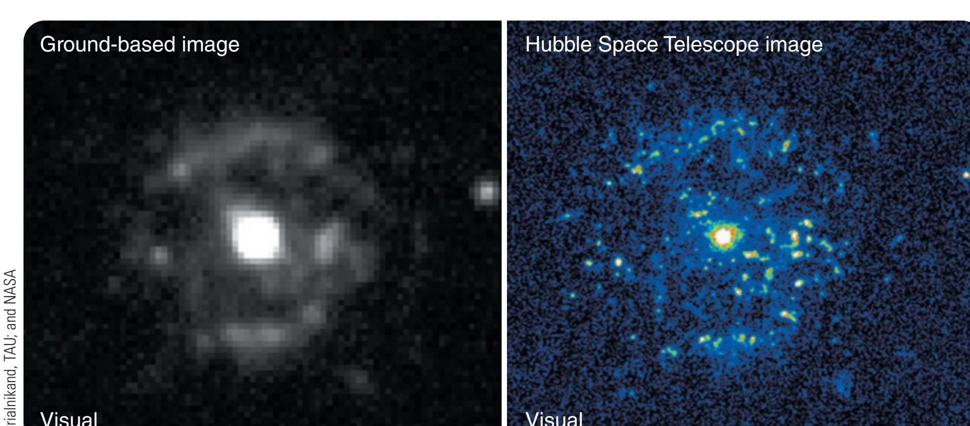
The End of Earth

Astronomy is about us. Although this chapter has discussed the deaths of stars, it has also been discussing the future of our planet. The sun is a medium-mass star and must eventually die by becoming a giant, possibly producing a planetary nebula, and collapsing into a white dwarf. That will spell the end of Earth.

Mathematical models of the sun suggest that it may survive for an additional 5 billion years or so, but it is already growing more

■ **Figure 15-13**

Nova T Pyxidis erupts about every two decades, expelling shells of gas into space. The shells of gas are visible from ground-based telescopes, but the Hubble Space Telescope reveals much more detail. The shell consists of knots of excited gas that presumably form when a new shell overtakes and collides with a previous shell.



luminous as it fuses hydrogen into helium. In a few billion years, it will exhaust hydrogen in its core and swell into a giant star about 100 times its present radius. That giant sun will be about as large as the orbit of Earth, so that will mark the end of our world. Whether the expanding sun becomes large enough to totally engulf Earth, its growing luminosity will certainly evaporate Earth's oceans, drive away the atmosphere, and even vaporize much of Earth's crust.

While it is a giant star, the sun will lose mass into space. This mass loss is a relatively gentle process, so any cinder that might remain of Earth would not be disturbed, although its orbit will grow larger as the sun's mass decreases. If anything is left of Earth after the sun becomes a giant, it will witness the sun's final collapse into a white dwarf. The atoms driven away from Earth will be part of the expanding nebula around the sun, and if it becomes hot enough it will ionize the expelled gas and light it up as a planetary nebula. Your atoms will be part of that nebula.

There is no danger that the sun will explode as a nova; it has no binary companion. And, as you will see, the sun is not massive enough to die the violent death of the massive stars.

The most important lesson of astronomy is that we are part of the universe and not just observers. The atoms we are made of are destined to return to the interstellar medium in just a few billion years. That's a long time, and it is possible that the human race will migrate to other planetary systems before then. That might save the human race, but our planet is stardust.

SCIENTIFIC ARGUMENT

How does spectroscopic evidence tell you what a nova explosion is like?

For this argument you need to use your knowledge of basic spectroscopy. As soon as a nova is seen, astronomers rush to telescopes to record spectra, and they see blueshifted absorption lines. The blueshifts are Doppler shifts showing that the near side of the object is coming toward Earth. The absorption lines must be formed by fairly dense gas seen through thinner gas much like the atmosphere of a star, so the surface of the star must be expanding rapidly outward. Later the spectrum becomes an emission spectrum, and Kirchhoff's laws tell you that the gas must have thinned. The continued blueshift shows that the expansion is continuing.

Now review the postexplosion evidence. **How do observations of novae long after they have faded provide evidence that white dwarfs are involved?**

15-4 The Deaths of Massive Stars

YOU HAVE SEEN that low- and medium-mass stars die relatively quietly as they exhaust their hydrogen and helium and then drive away their surface layers to form planetary nebulae. In contrast, massive stars live spectacular lives (■ Figure 15-14) and then destroy themselves in violent explosions.

Nuclear Fusion in Massive Stars

Stars on the upper main sequence have too much mass to die as white dwarfs, but their evolution begins much like that of their lower-mass cousins. They consume the hydrogen in their cores and ignite hydrogen shells; as a result, they expand into giants, or, for the most massive stars, supergiants. Next, their cores contract and fuse helium—first in the core and then in a shell, producing a carbon-oxygen core.

Unlike medium-mass stars, the massive stars do become hot enough to ignite carbon fusion at a temperature of about 1 billion Kelvin. Carbon fusion produces more oxygen and neon. As soon as the carbon is exhausted in the core, the core contracts, and carbon ignites in a shell. This pattern of core ignition and shell ignition continues with fuel after fuel, and the star develops the layered structure as shown in Figure 15-14, with a hydrogen-fusion shell above a helium-fusion shell above a carbon-fusion shell above . . . and so on. After carbon fuses, oxygen, neon, and magnesium fuse to make silicon and sulfur, and then the silicon fuses to make iron.

The fusion of these nuclear fuels goes faster and faster as the massive star evolves. Recall that massive stars must consume their fuels rapidly to support their great weight, but other factors also cause the heavier fuels like carbon, oxygen, and silicon to fuse at increasing speeds. For one thing, the amount of energy released per fusion reaction decreases as the mass of the fusing atom increases. To support its weight, a star must fuse oxygen much faster than it fused hydrogen. Also, there are fewer nuclei in the core of the star by the time heavy nuclei begin to fuse. Four hydrogens make a helium nucleus, and three heliums make a carbon, so there are 12 times fewer nuclei of carbon available for fusion than there were of hydrogen. This means the fusion of heavy elements goes very quickly in massive stars (■ Table 15-1). Hydrogen core fusion can last 7 million years in a 25-solar-mass star, but that same star will fuse the oxygen in its core in 6 months and its silicon in a day.

Supernova Explosions of Massive Stars

Theoretical models of evolving stars combined with nuclear physics allow astronomers to describe what happens inside a massive star when the last of its nuclear fuels are exhausted. The death of a massive star begins with iron nuclei and ends in cosmic violence.

Silicon fusion produces iron, the most tightly bound of all atomic nuclei (look back to Figure 7-11). Nuclear fusion is able to release energy by combining less tightly bound nuclei into a more tightly bound nucleus, but once the gas in the core of the star has been converted to iron, there are no nuclear reactions that can combine iron nuclei and release energy. The iron core is a dead end in the evolution of a massive star.

As a star develops an iron core, energy production begins to decline, and the core contracts. For nuclei less massive than iron, such contraction heats the gas and ignites new fusion fuels, but nuclear reactions involving iron remove energy from

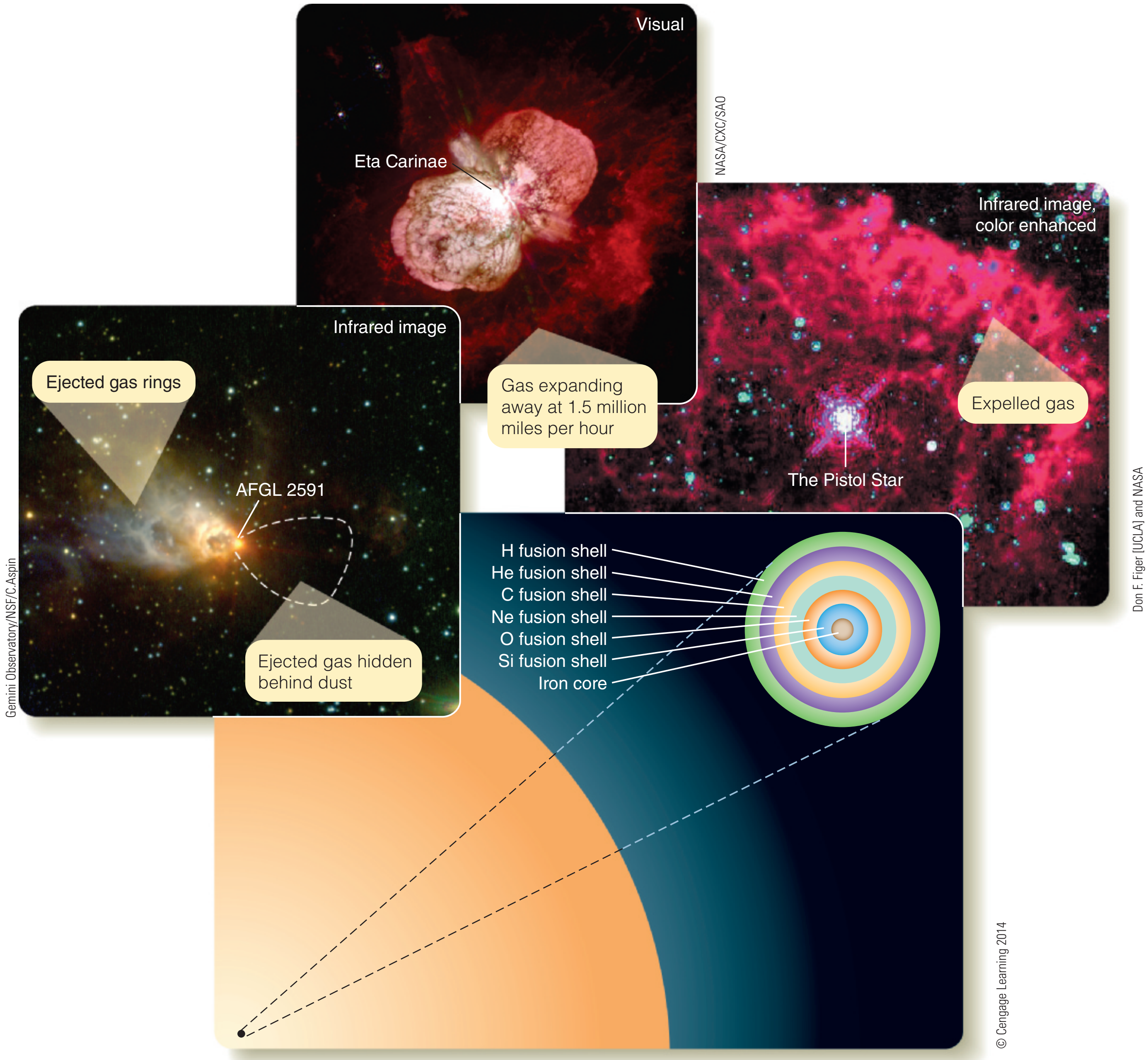


Figure 15-14

Massive stars live fast and die young. The three shown here are among the most massive stars known, containing 100 solar masses or more. They are rapidly ejecting gas into space. The centers of these massive stars develop Earth-size cores (magnified 100,000 times in this figure) composed of concentric layers of gases undergoing nuclear fusion. The iron core at the center leads eventually to a star-destroying explosion.

the core in two ways. First, the iron nuclei begin capturing electrons and breaking into smaller nuclei. The gas is so dense it is degenerate, and the degenerate electrons helped support the core. The loss of some of the electrons allows the core to contract even faster. Second, temperatures are so high that the average photon is a high-energy gamma ray, and these gamma rays are absorbed by atomic nuclei, causing them to break into smaller fragments. The removal of the gamma rays also allows the core to contract even faster. Although the core of the star can't generate energy by nuclear fusion, it can draw on the tremendous energy stored in its gravitational field. As the core contracts, the temperature shoots up, but it is not enough to stop the contraction, and the core of the star collapses inward in less than a tenth of a second.

Table 15-1 | Heavy-Element Fusion in a 25-$M_{\odot}$ Star

Fuel	Time	Percentage of Lifetime
H	7,000,000 years	93.3
He	500,000 years	6.7
C	600 years	0.008
O	0.5 years	0.000007
Si	1 day	0.00000004

This collapse happens so rapidly that the most powerful computers are unable to predict the details. Consequently models of supernova explosions contain approximations. Nevertheless, the models predict that the collapse of the core produces an immense supernova explosion in which the star's outer layers are blasted outward. The core of the star must quickly become a neutron star or a black hole, the subjects of the next chapter.

To understand how the inward collapse of the core can produce an outward explosion, it helps to think about a traffic jam. The collapse of the innermost part of the degenerate core allows the rest of the core to fall inward creating a tremendous traffic jam as all of the nuclei fall toward the center. It is as if every car owner in Indiana suddenly tried to drive into downtown Indianapolis. There would be a traffic jam not only downtown but also in the suburbs; and, as more cars arrived, the traffic jam would spread outward. Similarly, although the innermost core collapses inward, a shock wave (a "traffic jam") develops and moves outward through the rest of the star.

The shock wave moves outward through the star aided by two additional sources of energy. First, when the iron nuclei in the core are disrupted, they produce a flood of neutrinos. In fact, for a short time the collapsing core produces more energy per second than all of the stars in all of the visible galaxies in the universe, and 99 percent of that energy is in the form of neutrinos. This flood of neutrinos carries large amounts of energy out of the core, allowing it to collapse further, helps heat the gas outside the core, and accelerates the outward-bound shock wave. The torrent of energy flowing out of the core also triggers tremendous turbulence, and intensely hot gas rushes outward from the interior (Figure 15-15). Again, this rising hot gas carries energy out into the envelope and helps drive the shock wave outward. Within a few hours, the shock wave bursts outward through the surface of the star and blasts it apart.

The supernova seen from Earth is the brightening of the star as its distended envelope is blasted outward by the shock wave.

The Exploding Core of a Supernova

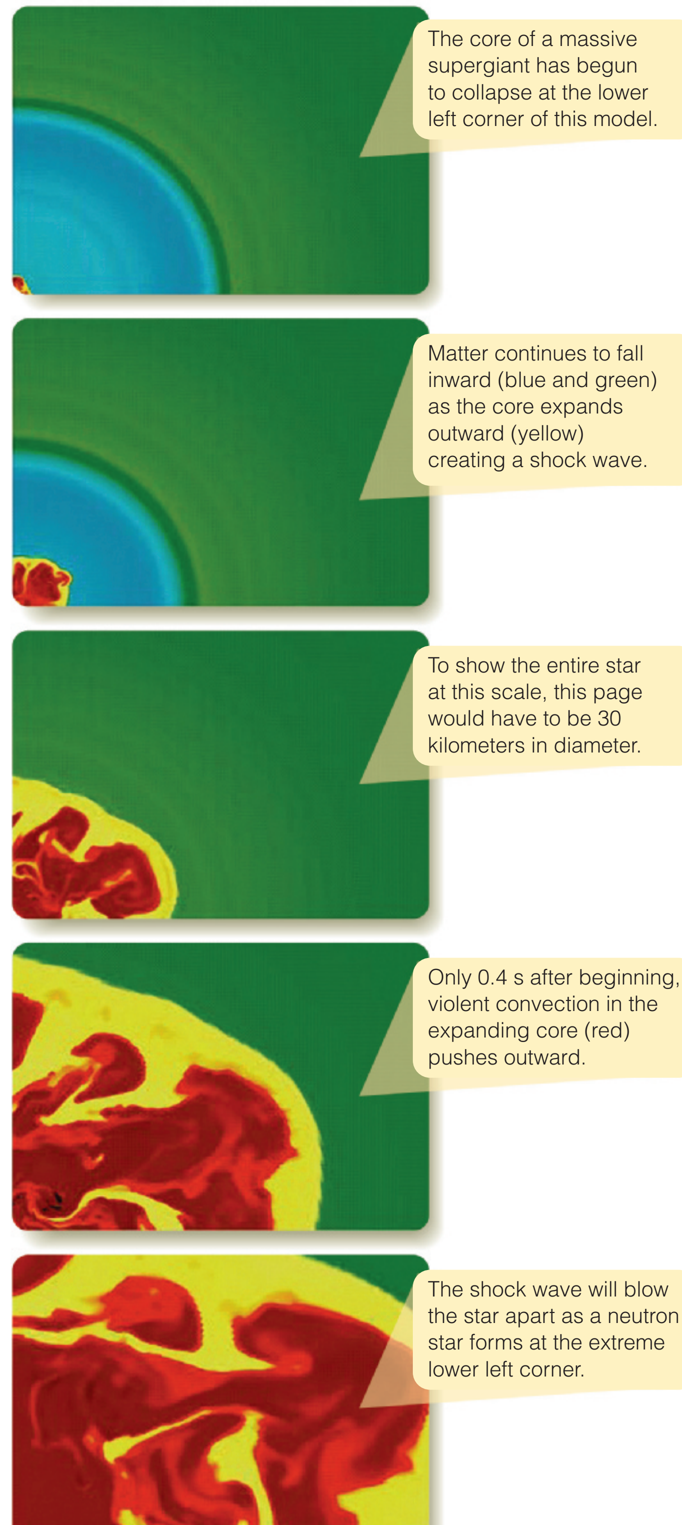


Figure 15-15

As the iron core of a massive star begins to collapse, intensely hot gas triggers violent convection. Even as the outer parts of the core continue to fall inward, the turbulence blasts outward and reaches the surface of the star within hours, creating a supernova eruption. This diagram is based on mathematical models and shows only the exploding core of the star.

As months pass, the cloud of gas expands, thins, and fades, and the rate at which it fades matches the decay rate of certain radioactive nuclei produced in the explosion. The violence in the outer layers can create densities and temperatures high enough to trigger nuclear fusion reactions that produce as much as half a solar mass of radioactive nickel-56. The nickel gradually decays to form radioactive cobalt, which decays to form normal iron. Essentially all of the iron in the core of the star is destroyed when the core collapses, but more iron is produced in the outer layers and that releases energy that keeps the supernova glowing.

The presence of nuclear fusion in the outer layers of the supernova testifies to the violence of the explosion. A typical supernova is equivalent to the explosion of 10^{28} megatons of TNT—about 3 million solar masses of high explosive. But, of course, the explosion is entirely silent. It is a **Common Misconception** promoted by science fiction movies and television that explosions in space are accompanied by sound. You know that's not true. Space is nearly a perfect vacuum, and sound can't travel through a vacuum. Supernova explosions are among the most violent events in nature, but they are silent.

Collapsing massive stars can trigger violent supernova explosions. There is, however, more than one kind of supernova.

Classifying Supernovae

A supernova occurs when a star with a degenerate core collapses, but astronomers have found different kinds of supernovae. Those that do not have hydrogen lines in their spectra are called **type I supernovae,** and those that do have hydrogen lines are **type II supernovae.** As you read in the previous section, a massive star with a degenerate core can collapse and produce a supernova. These are classified type II and contain hydrogen lines in their spectra because those massive stars contain lots of hydrogen. But there are other ways a star with a degenerate interior can collapse.

A supernova can occur if a white dwarf in a binary system gains mass from its companion, exceeds the Chandrasekhar limit, and collapses. The interior of the white dwarf is degenerate, and as the collapse begins, temperature and density increase and the carbon–oxygen core begins to fuse. Because the gas is degenerate there is no thermostat to regulate the fusion, and the entire white dwarf fuses suddenly in a violent explosion classified as a **type Ia supernova.** The white dwarf is entirely destroyed, and there are no hydrogen lines in the spectrum of the supernova because white dwarfs contain little hydrogen.

The much less common **type Ib supernova** can occur if a massive star loses its hydrogen-rich outer layers to a companion star. The remains of the massive star can develop a degenerate core and collapse, as described in the previous section, producing a supernova explosion that lacks hydrogen lines. A type Ib supernova is just a type II supernova in which the star has lost its outer layers.

As you can see, all supernovae involve the collapse of a star with a degenerate interior in which there is no thermostat to keep the reactions under control. Modern computers allow astronomers to study the interiors of exploding stars, but the companion to theory is observation, so you should ask what observational evidence supports the story of supernovae explosions.

A History of Supernovae

Because supernovae are so rare they are difficult to observe, astronomers must depend on historical records. Some records go back centuries, but more recent supernovae give astronomers the best evidence against which to test their theories.

In the year 1054, Chinese astronomers saw a "guest star" appear in the constellation we know as Taurus. It became so bright it was visible in the daytime for a month and took nearly two years to fade from sight. At the site of that ancient supernova, modern telescopes reveal a many-legged nebula known as the Crab Nebula (■ Figure 15-16). Doppler shifts show that the nebula is expanding at 1400 km/s, and it is now about 1.4 pc in radius. Dividing distance by velocity reveals that the nebula began expanding nine or ten centuries ago, about when the "guest star" was seen.

The hazy glow inside the expanding filaments of the Crab Nebula is **synchrotron radiation** that is produced by high-speed electrons spiraling through a magnetic field. In the Crab Nebula, the electrons are moving so fast they produce visible light. By now those high-speed electrons should have slowed down, but astronomers have evidence that the electrons continue to be energized by a neutron star left behind by the supernova. You will learn more about neutron stars in the next chapter.

Only a few supernovae have been visible to the naked eye. Arab astronomers saw one in the year 1006, and the Chinese saw one in 1054. Tycho's supernova appeared in 1572 and Kepler's supernova in 1604. Also, the guest stars of the years 185, 386, 393, and 1181 may have been supernovae.

Most supernovae are seen in distant galaxies (■ Figure 15-17), and modern observations show that type I and type II explosions can be distinguished by the shapes of their light curves. Type I supernovae are about six times more luminous at maximum than type II, and they fade more slowly at first (■ Figure 15-18).

Since the invention of the telescope in 1609, no one had seen a bright supernova until 1987 when astronomers in Chile spotted a naked-eye supernova brightening in the southern sky (■ Figure 15-19). Known as Supernova 1987A (SN 1987A), it exploded only 53,000 pc away in the Large Magellanic cloud, a small nearby galaxy. The spectrum is that of a type II supernova, but its light curve was not typical (Figure 15-18). It was evidently produced by the collapse of a hot, blue supergiant rather than the usual cool, red supergiant. The star had been a red supergiant a few thousand years ago but contracted and heated before it exploded. Because it was smaller, it did not reach maximum light as quickly as most.

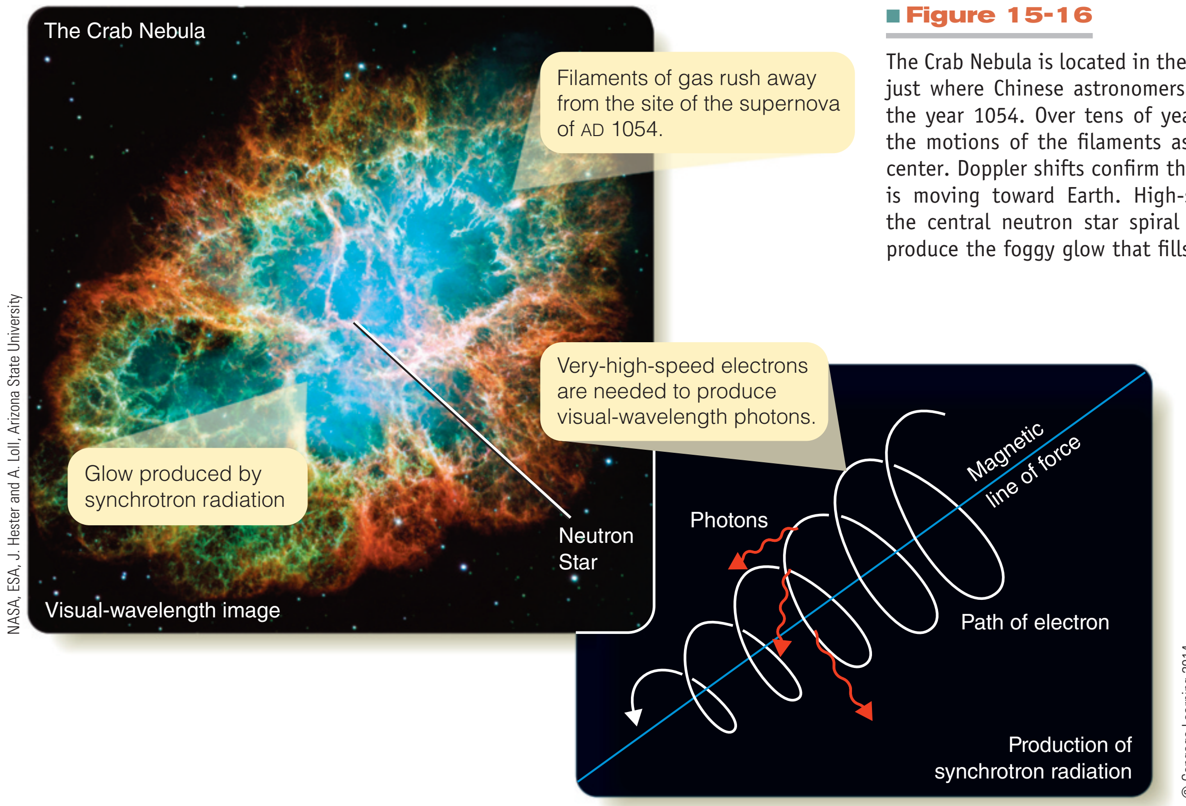


Figure 15-16

The Crab Nebula is located in the constellation Taurus the Bull, just where Chinese astronomers saw a brilliant guest star in the year 1054. Over tens of years, astronomers can measure the motions of the filaments as they expand away from the center. Doppler shifts confirm that the near side of the nebula is moving toward Earth. High-speed electrons produced by the central neutron star spiral through magnetic fields and produce the foggy glow that fills the nebula.

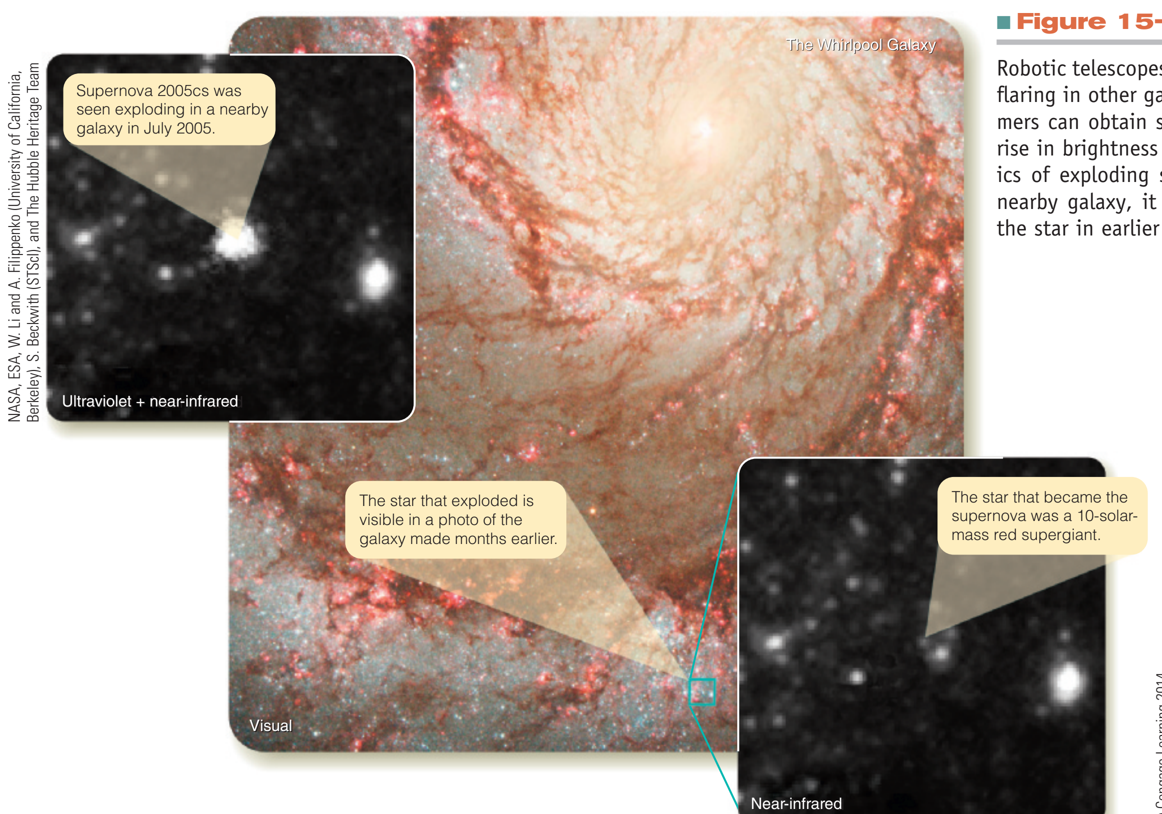
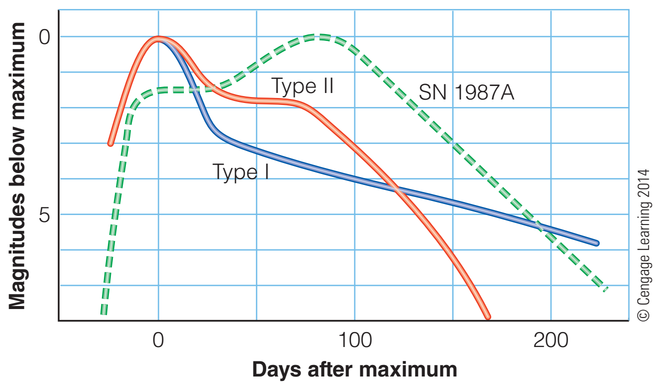


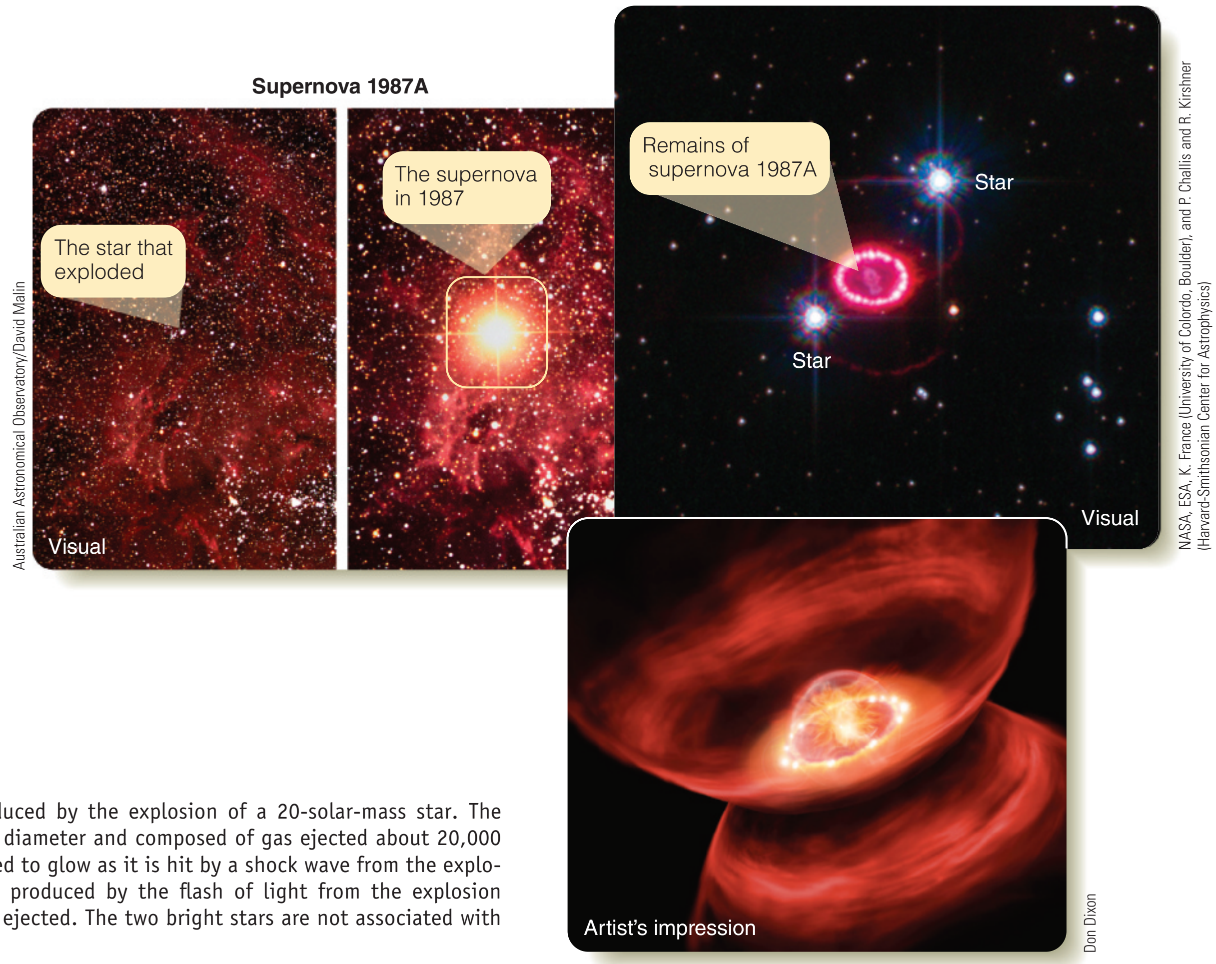
Figure 15-17

Robotic telescopes search every night for supernovae flaring in other galaxies. When one is seen, astronomers can obtain spectra and record the supernova's rise in brightness and its decline to study the physics of exploding stars. If a supernova is seen in a nearby galaxy, it is sometimes possible to identify the star in earlier photos.



■ Figure 15-18

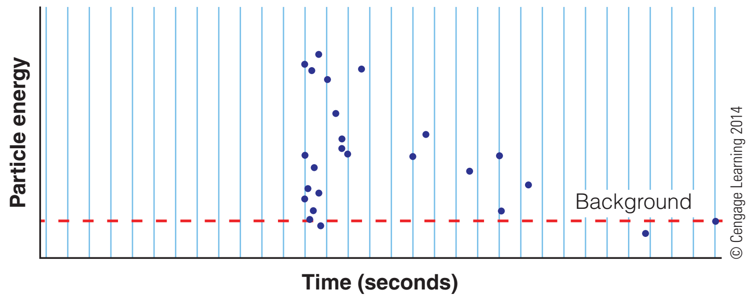
Type I supernovae decline rapidly at first and then more slowly, but type II supernovae pause for about 100 days before beginning a steep decline. Supernova 1987A was odd in that it did not rise directly to maximum brightness. These light curves have been adjusted to the same maximum brightness.



■ Figure 15-19

Supernova 1987A was produced by the explosion of a 20-solar-mass star. The bright ring is about 1 ly in diameter and composed of gas ejected about 20,000 years ago. It is being excited to glow as it is hit by a shock wave from the explosion. The fainter rings are produced by the flash of light from the explosion illuminating gas previously ejected. The two bright stars are not associated with the supernova.

Supernova 1987A was later located on a survey photograph made 19 hours before the official discovery, but a critical observation was made even before that. About 3 hours before the supernova was first noticeable in photographs, instruments on Earth detected 25 neutrinos arriving from the direction of 1987A during a span of just 13 seconds (■ Figure 15-20). Neutrinos are so difficult to detect that those few mean that some 10^{17} must have swept through the detectors in those seconds. The blast was so intense that in a few seconds roughly 20 trillion neutrinos passed harmlessly through each human being on Earth, 170,000 light years away. Evidently the burst of neutrinos was released when the iron core collapsed into a neutron star, and the supernova itself was first detected about three hours later when the shock wave blasted the star's surface into space.



■ **Figure 15-20**

About 3 hours before the supernova was detectable in photographs and 22 hours before it was first seen at visual wavelengths, detectors on Earth recorded 25 neutrinos arriving from the direction of the supernova. The burst dramatically exceeded the background of low-energy, sporadic neutrinos normally detected.

New observations suggest other ways stars can explode. Six supernovae have been observed that are unusually luminous but emit much of their light in the ultraviolet. Astronomers wonder if they are produced by the collapse of star that is spinning rapidly or a star that has recently ejected a shell of hydrogen poor gas. They may represent a new class of supernovae.

Supernova Remnants

Supernovae typically leave behind shells of hot gas expanding outward at velocities as high as 20,000 km/s and carrying away a fifth of the mass of the star. The shells can sweep up more gas from the interstellar medium and heat it to produce a **supernova remnant**, the nebular remains of a supernova explosion (■ Figure 15-21).

Some supernova remnants, such as Cassiopeia A, show evidence of jets of matter rushing outward in opposite directions. These may have been ejected as the rotating star collapsed, and, conserving angular momentum, spun up to very high speeds. The first matter blown outward from such a rapidly rotating star could have emerged as jets from its poles. Astronomers are just beginning to understand the details of such violent explosions.

Supernova remnants look quite delicate and do not survive very long—a few tens of thousands of years—before they gradually mix with the interstellar medium and vanish. The Crab Nebula is a young remnant, only about 950 years old and about 8.8 ly in diameter. Older remnants can be larger. Some supernova remnants are visible only at radio and X-ray wavelengths. They have become too tenuous to emit detectable light, but the collision of the expanding hot gas with the interstellar medium can generate radio and X-ray radiation. You saw in the previous chapter that the compression of the interstellar medium by expanding supernova remnants can also trigger star formation.

SCIENTIFIC ARGUMENT

What evidence do astronomers have that Supernova 1987A formed a neutron star?

The critical observation consists of only 19 neutrinos detected coming from the direction of the supernova. Because neutrinos are so difficult to detect, those 19 must mean that a huge flood of neutrinos passed through Earth just before the supernova was seen. Theory says the collapse of the core into a ball of neutrons should release a tremendous burst of neutrinos, and astronomers link the neutrinos that were detected with the formation of that ball of neutrons. Notice that this evidence depends on a theory. That's not unusual, but scientists are very careful in analyzing such evidence to be sure the background theory is right. Only then is the evidence meaningful.

Now create a new argument. **What evidence can you cite that type Ia supernovae are not produced by massive stars?**

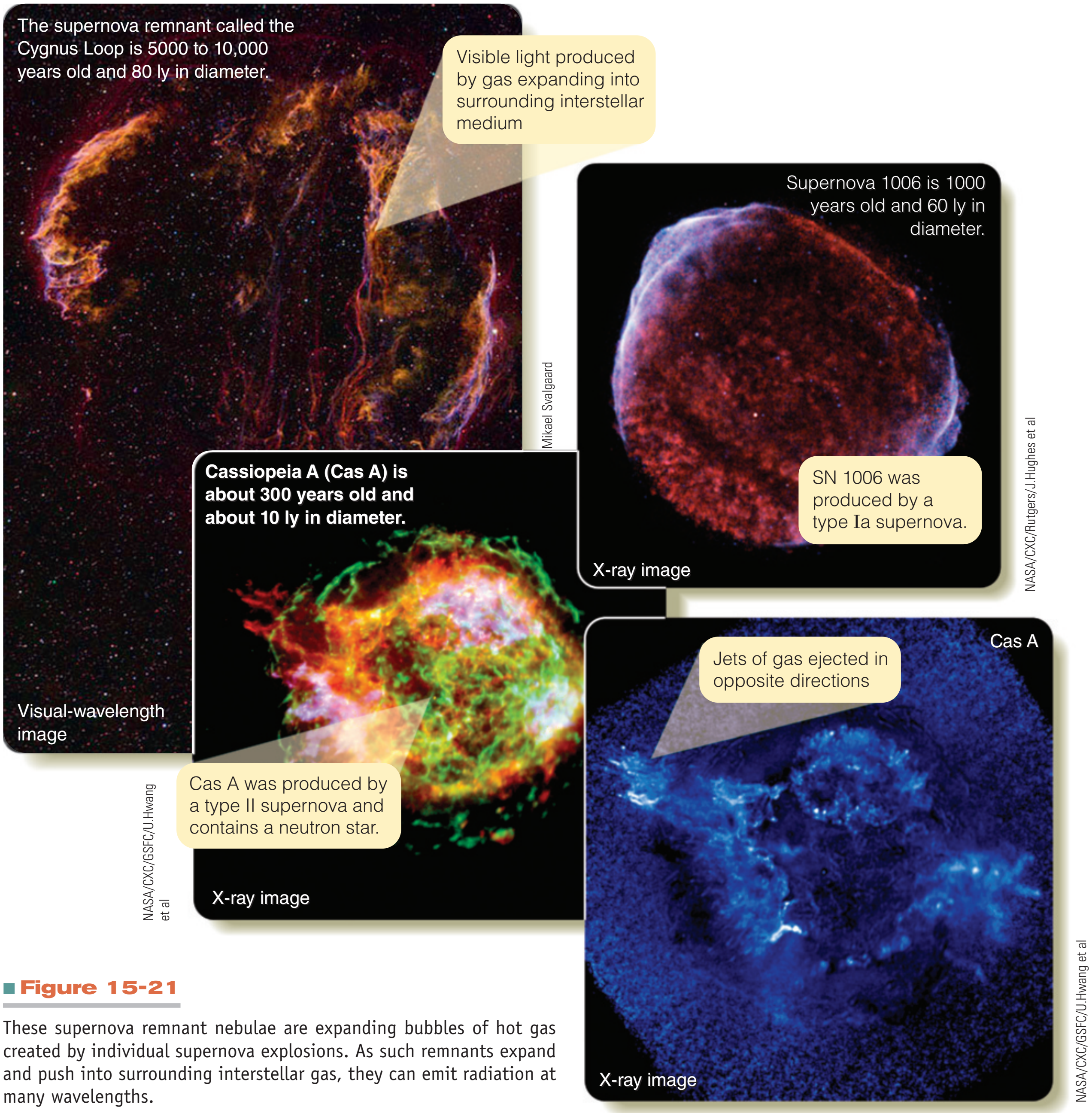


Figure 15-21

These supernova remnant nebulae are expanding bubbles of hot gas created by individual supernova explosions. As such remnants expand and push into surrounding interstellar gas, they can emit radiation at many wavelengths.

What Are We? Stardust

You are made of atoms that were cooked up inside stars. Gravity draws matter together to make stars, and, although nuclear fusion delays gravity's final victory, stars must eventually die. That process of star life and star death manufactures atoms heavier than helium and spreads them back into the interstellar medium where they can become part of the gas clouds that form new stars. All of the atoms in your body except for the hydrogen were made inside stars.

Some of your atoms such as the carbon were cooked up in the cores of medium-mass stars like the sun and were puffed out into space when those stars died and produced planetary nebulae. Some of your atoms, such as the calcium in your bones, were made inside massive stars and were blown out into space during type II supernova explosion. Many of the iron atoms in your blood were made by the sudden fusing of carbon when white dwarfs collapsed in type Ia supernova explosions. In fact, a few of your heavier atoms such as iodine in your thyroid gland and selenium in your nerve cells were produced in the raging violence of supernova explosions.

You are made of star stuff scattered into space long ago by the violent deaths of stars. What are we? We are stardust.

Study and Review

Summary

- A **nova (p. 318)** appears to be a new star that becomes visible and then fades after a few weeks. **Supernovae (p. 318)** are more luminous and last longer. Both are associated with the deaths of stars.
- Main sequence stars like the sun can expand and become giant stars when they use of the hydrogen fuel in their cores. As that happens, the core contracts and heats up. Hydrogen fusion begins in a spherical layer around the core—a hydrogen-fusion shell.
- Energy from the hydrogen-fusion shell swells the star into a cool giant 10 to 100 times larger in diameter than the sun. Massive stars swell to become supergiants up to 1000 times larger in diameter than the sun.
- The contraction of the star's core eventually ignites helium fusion, first in the core and later in a shell. Fusion of three helium nuclei into one carbon nucleus is called the **triple-alpha process (p. 320).** If the star is massive enough, it can eventually fuse carbon and other elements.
- In **degenerate matter (p. 320),** the density is so high quantum mechanical effects prevent electrons from changing their energies. Such matter is very difficult to compress, and its pressure does not depend on its temperature.
- If a star's mass lies between about 0.4 and 3 solar masses, its helium core becomes degenerate before the helium ignites. Because pressure does not depend on temperature, there is no pressure–temperature thermostat to control the reactions, and when helium fusion ignites the core explodes in a **helium flash (p. 321).** All of the energy produced is absorbed by the star.
- As the giant star fuses helium in its core and hydrogen in a shell, it moves toward the hot side of the H–R diagram. As soon as it exhausts the helium in its core and begins fusing helium in a shell, it moves back toward the cool side producing a loop in its evolutionary path.
- You can see evidence of stellar evolution in the H–R diagrams of star clusters. Stars in both **open clusters (p. 324)** and in **globular clusters (p. 324)** evolve in similar ways. The stars begin their evolution at the same time but evolve at different rates, depending on their masses. The most massive stars leave the main sequence first and are followed later by progressively less massive stars. This makes the evolution of stars visible in the H–R diagram.
- You can estimate the age of a star cluster from the **turnoff point (p. 324)** in its H–R diagram.
- In old clusters, stars fusing helium follow a loop in the H–R diagram that is visible in the diagrams of star clusters as the **horizontal branch (p. 325).**
- Red dwarfs less massive than about 0.4 solar mass are completely mixed and will have very little hydrogen left when they die. They can't ignite a hydrogen-fusion shell, so they can't become giant stars. They will remain on the main sequence for many times the present age of the universe.
- Medium-mass stars like the sun become cool giants and fuse helium but can't fuse carbon. They eventually blow away their outer layers and collapse into hot white dwarfs. Ultraviolet radiation from the white dwarfs ionize the gas to produce **planetary nebulae (p. 326).**
- White dwarfs can't contract as they cool and will eventually become **black dwarfs (p. 327).** The **Chandrasekhar limit (p. 327)** shows that no white dwarf more massive than 1.4 solar masses can be stable. Presumably more massive stars can become white dwarfs only if they shed mass.
- Close binary stars evolve in complex ways because they can transfer mass from one star to the other. The two stars are enclosed by a dumbbell region known as the **Roche lobes (p. 331).** If mass from one star crosses the surface of these lobes, the **Roche surface (p. 331),** it can fall into the other star. The **Lagrange points (p. 331)** in the rotating system are points where mass can remain stable. Matter can flow between the stars through the **inner Lagrange point (p. 331),** which connects the two Roche lobes.
- As a star becomes a giant, it can expand and fill its Roche lobe, spilling mass to the other star. Mass transfer explains why some binary systems contain a main-sequence star more massive than its giant companion—the Algol paradox.
- Mass that is transferred from one star to the other must conserve **angular momentum (p. 332)** and can form a whirling **accretion disk (p. 332)** around the receiving star. Accretion disks can become hot enough to emit light and even X-rays.
- Mass transferred onto the surface of a white dwarf can build up a layer of fuel that erupts in a nova explosion. A white dwarf can erupt repeatedly so long as mass transfer continues to form new layers of fuel.
- The evolution of the sun into a giant and then its collapse into a white dwarf will end life on Earth.
- Stars more massive than about 8 solar masses can't lose mass fast enough to reduce their mass low enough to die by ejecting a planetary nebula and collapsing into a white dwarf. Such massive stars must die more violent deaths.
- The massive stars on the upper main sequence fuse nuclear fuels one after the other producing a layering of fusion shells, but such stars can't fuse iron because iron is the most tightly bound of all atomic nuclei. When an aging massive star forms an iron core, the core collapses and triggers a supernova explosion known as a **type II supernova (p. 337).**
- The spectra of type II supernovae contain hydrogen lines, but the spectra of **type I supernovae (p. 337)** do not. At least two causes of type I supernovae are known.
- A **type Ia supernova (p. 337)** can occur when mass transferred onto a white dwarf pushes it over the Chandrasekhar limit and it collapses suddenly, fusing all of its carbon at once. A **type Ib supernova (p. 337)** occurs when a massive star in a binary system loses its outer layers of hydrogen before it explodes.
- A supernova expels an expanding shell of gas called a **supernova remnant (p. 340).** The supernova seen in the year 1054 produced a supernova remnant known as the Crab Nebula, which emits **synchrotron radiation (p. 337),** evidence of a powerful energy source remaining inside the remnant.
- The supernova 1987A is only a few years old, but its expanding gases will eventually form a supernova remnant. Neutrinos observed coming from the direction of the supernova are evidence that the core collapsed and formed a neutron star.

Study and Review

Review Questions

1. Why does helium fusion require a higher temperature than hydrogen fusion?
2. How can the contraction of an inert helium core trigger the ignition of a hydrogen-fusion shell?
3. Why does the expansion of a star's envelope make it cooler and more luminous?
4. Why is degenerate matter so difficult to compress?
5. How does the presence of degenerate matter in a star trigger the helium flash?
6. How can star clusters confirm astronomers' theories of stellar evolution?
7. Why don't red dwarfs become giant stars?
8. What causes an aging giant star to produce a planetary nebula?
9. Why can't a white dwarf contract as it cools? What is its fate?
10. Why can't a white dwarf have a mass greater than 1.4 solar masses?
11. How can a star of as much as 8 solar masses form a white dwarf when it dies?
12. How can you explain the Algol paradox?
13. How can the inward collapse of the core of a massive star produce an outward explosion?
14. What is the difference between type I and type II supernovae?
15. What is the difference between a supernova explosion and a nova explosion?
16. **How Do We Know?** In what ways do the appearance of supernova explosions depend on the properties of subatomic particles?

Discussion Questions

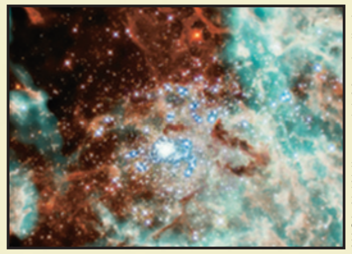
1. How do you know the helium flash occurs if it can't be observed? Can you accept something as real if you can never observe it?
2. False-color radio images and time-exposure photographs of astronomical images show aspects of nature you can never see with unaided eyes. Can you think of common images in newspapers or on television that reveal phenomena you can't see?

Problems

1. About how long will a 0.4-solar-mass star spend on the main sequence?
2. If the stars at the turnoff point in a star cluster have masses of about 4 solar masses, how old is the cluster?
3. About how far apart are the stars in an open cluster? In a globular cluster? (*Hints:* What share of the cluster's volume belongs to a single star? *Note:* The volume of a sphere is $\frac{4}{3}\pi r^3$.)
4. The Ring Nebula in Lyrae is a planetary nebula with an angular diameter of 76 arc seconds and a distance of 5000 ly. What is its linear diameter? (*Hint:* Use the small-angle formula, Chapter 3.)
5. Suppose a planetary nebula is 1 pc in radius. If the Doppler shifts in its spectrum show it is expanding at 30 km/s, how old is it? (*Note:* 1 pc equals 3.1×10^{13} km, and 1 year equals 3.2×10^7 seconds, to 2 significant figures.)
6. If a star the size of the sun expands to form a giant 20 times larger in radius, by what factor will its average density decrease? (*Note:* The volume of a sphere is $\frac{4}{3}\pi r^3$.)
7. If a star the size of the sun collapses to form a white dwarf the size of Earth, by what factor will its density increase? (*Note:* The volume of a sphere is $\frac{4}{3}\pi r^3$. See Appendix A for the radii of the sun and Earth.)
8. The Crab Nebula is now 1.35 pc in radius and is expanding at 1400 km/s. About when did the supernova occur? (*Note:* 1 pc equals 3.1×10^{13} km.)
9. If the Cygnus Loop is 25 pc in diameter and is 10,000 years old, with what average velocity has it been expanding? (*Note:* 1 pc equals 3.1×10^{13} km, and 1 year equals 3.2×10^7 seconds, to 2 significant figures.)
10. Observations show that the gas ejected from SN 1987A is moving at about 10,000 km/s. How long will it take to travel one astronomical unit? One parsec? (*Note:* 1 AU equals 1.5×10^8 km, and 1 pc equals 3.1×10^{13} km, to 2 significant figures.)

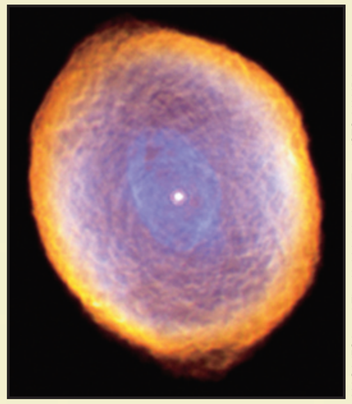
Learning to Look

1. The star cluster in the photo at the right contains many hot, blue, luminous stars. Sketch its H–R diagram and discuss its probable age.



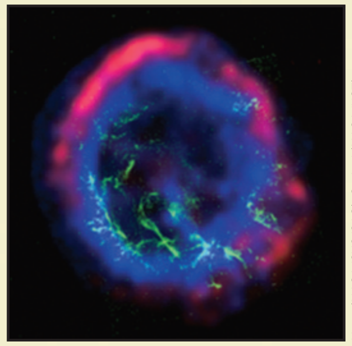
NASA, N. Walborn and J. Maíz-Apellániz (Space Telescope Science Institute, Baltimore, MD), R. Barbá (La Plata Observatory, La Plata, Argentina)

2. What processes caused a medium-mass star to produce the nebula at the right? The nebula is now about 0.1 ly in diameter and still expanding. What will happen to it?



NASA/ESA and The Hubble Heritage Team/STScI/AURA

3. The image at right combines X-ray (blue), visible (green), and radio (red) images. Observations show the sphere is expanding at a high speed and is filled with very hot gas. What kind of object produced this nebula? Roughly how old do you think it must be?



X-ray (NASA/CXC/SAO); Optical (NASA/HST); Radio (ACTA)

Great Debates

1. *Fate of Humans?* The sun will eventually go through its red giant stage, and during that time it may expand to the orbit of Earth. Will Earth survive this stage of the sun's evolution?
   a. Use at least three vocabulary words from the textbook correctly in your debate, including the stages of the sun's evolution; underline each; and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.
2. *Can Humans Be Considered Degenerate?* Given the definition provided in your textbook of "degeneracy" for electrons and neutrons, do you think humans can be considered degenerate?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.
3. *M Dwarf Evolution.* M dwarf stars on the main sequence have between 0.08 and 0.4 solar masses. These very-low-mass stars live very long lives. For example, a 0.1 solar mass star can live as long as 6 trillion years, 400 times longer than the current age of the universe. Given that no M dwarf stars have yet died off, what do you think an M dwarf star will die into? Will it die into a neutron star, black hole star, a white dwarf star, or none of the above?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.
4. *Neutrino Burning?* In the movie *2012*, neutrinos interact with matter. For example, neutrinos heat up Earth's inner core, making the core boil. The boiling core transports energy to Earth's surface, causing buckling of the crust and the building of mountains and valleys. Is any of this possible? Do neutrinos interact with matter?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources that support your stand.
   c. Cite your sources.
5. *Another Heavy Bombardment Period?* When the sun goes through its red giant stage, the planets are expected to migrate outward from their current location, stirring up the asteroid and Kuiper belts. These stirrings might result in comet and asteroid impacts on the planets and their satellites, including the moon and Earth. It is currently speculated that about one-third of white dwarf systems have these debris disks of comet and asteroids. Do you think Earth will experience another heavy bombardment period? Will the moon protect the Earth from the infall? Should humans be concerned about another heavy bombardment period, and should humans do anything to prepare?
   a. Use at least three vocabulary words from the textbook correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

ENHANCED WebAssign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Mass-Star Lifetime Relation
- Post Main-Sequence Evolution
- Cluster Turnoff
- Roche Lobes
- Nova Mechanism
- Inside Stars

CengageNOW Virtual Astronomy Labs 2.0

Virtual ASTRONOMY LABORATORIES

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 11: Spectral Sequence

The H–R diagram, invented independently in the early twentieth century by Netherlands astronomer Ejnar (pronounced *EYE-nar*) Hertzsprung and U.S. astronomer Henry Norris Russell, is the basic tool that astronomers use to this day to understand stars, star clusters, and even galaxies consisting of billions of stars. By plotting stellar absolute magnitude (or, equivalently, luminosity) on the vertical axis versus stellar spectral type (or, equivalently, color or surface temperature) on the horizontal axis, Hertzsprung and Russell found that stars commonly have only some combinations of those properties. The great majority of stars, including the sun, are plotted along the so-called main sequence that extends diagonally from the luminous/hot/blue/O-type upper left corner to the faint/cool/red/M-type lower right corner. Only a few stars have properties in other parts of the diagram.

The H–R diagram is directly about the "outsides" of stars (their photospheres and atmospheres). But the properties of stellar exteriors must be determined by what's going on in their interiors. From its invention, the H–R diagram was considered also to be an evolutionary diagram. Astronomers made guesses about which positions on the H–R diagram represented young stars and which represented old stars. The true situation took many decades for full understanding.

By making stellar interior model calculations based on positions of stars in the H-R diagram, astronomers now know that protostars start far off to the right (with low temperatures) and then contract and heat up to become main-sequence stars. As you learned in thus chapter, once in its "home" position on the main sequence determined by its mass, a star's properties stay almost constant for a relatively long main-sequence lifetime, with a gradual rise in luminosity as it ages (H–R diagram point slowly rising). Then, when a star begins to run out of hydrogen for fusion fuel, it leaves the main sequence in a brief interval (astronomically speaking), and spends the remainder of its life moving around in the red giant "corner." The more massive stars, using their fuel supply disproportionately quickly, leave the main sequence sooner than less massive stars.

Section 3 of Virtual Astronomy Lab 11, "Spectral Sequence," first lets you use the principles of stellar evolution to study the behavior of some idealized stars. Then, comparing the properties of those stars with the position of a star cluster's H–R diagram main-sequence "turn-off point," where stars are about to leave the main sequence and become red giants, gives you the age of the cluster as a whole. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 13: Stellar Explosions, Novae, and Supernovae

You have never touched degenerate matter. It forms under only the most extreme conditions, such as those inside a white dwarf stellar remnant, where the mass of a star is compressed into a volume the size of Earth and atoms of elements such as helium, carbon, and oxygen are subjected to immense pressures. The nuclei are still identifiably helium or carbon or oxygen, but the material is in an exotic state with strange properties that are not found on Earth.

If you could orbit a white dwarf and drop matter onto its surface, you would discover that you can't make it bigger. Its radius actually shrinks as you add matter. In fact, according to the equations that describe degenerate matter, if you add enough matter, the radius shrinks to zero. What the equations are really telling you is that a white dwarf cannot be stable above a certain mass. If its mass is too large, it must collapse and transform into an even stranger object, either a neutron star or a black hole.

Another important property of degenerate matter is that its pressure does not depend on its temperature. That means the natural pressure–temperature thermostat that (as you learned in the previous chapter) controls nuclear reactions inside normal stars doesn't work in an object made of degenerate matter. There is no problem if there are no nuclear reactions, but there are two circumstances that can cause trouble. If a layer of matter accumulates on the surface of a white dwarf and becomes hot enough, it can start fusion there without a controlling thermostat, resulting in a nova explosion. Furthermore, if a white dwarf gains enough matter, the helium, carbon, or oxygen in the interior can begin fusion. Because the pressure–temperature thermostat can't operate in degenerate matter, the entire white dwarf can detonate in a type Ia supernova.

These violent explosions may sound like bad news, but they do something that is important to you personally. In the moments of the explosions, nova or supernova, temperatures can shoot so high that heavy elements can be created. Some of the atoms playing vital roles in your body's metabolism, such as iodine in your thyroid gland and selenium in your nerves, were made inside exploding stars long ago. Also silver, gold, and platinum are rare valuable elements because they are made only in supernova explosions.

Virtual Astronomy Lab 13, "Stellar Explosions, Novae, and Supernovae," first gives you an opportunity to compute densities of white dwarf stars and explore the consequences of degenerate matter's properties. You will also learn how important type Ia supernovae are as distance indicators. Sign in at http:// login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

16

Neutron Stars and Black Holes

Guidepost

In the last two chapters you have traced the story of stars from birth to death. By now you are asking a simple question, "What's left?" The answer depends on the mass of the star. You already know that stars, like the sun, produce white dwarf corpses, but more massive stars leave behind the strangest beasts in the cosmic zoo.

Now you are ready to meet neutron stars and black holes, and your exploration will answer four essential questions:

- **How does theory predict the existence of neutron stars?**
- **How do astronomers know neutron stars really exist?**
- **How does theory predict the existence of black holes?**
- **How can astronomers be sure that black holes really exist?**

This chapter will show you more striking examples of how astronomers combine observations and theory to understand nature.

This chapter ends the story of individual stars, but it does not end the story of stars. In the next chapter, you will begin exploring the giant communities in which stars live—the galaxies.

In this artist's impression, matter from a normal star flows into an accretion disk around a compact object. Processes in the spinning disk eject gas and radiation in jets perpendicular to the disk.

Copyright © 2010 by Fahad Sulehria, www.novacelestia.com

Almost anything is easier to get into than out of.

AGNES ALLEN

GRAVITY ALWAYS WINS. No matter how long a star struggles to withstand its own gravity, it must eventually exhaust its fuels and die by collapsing. Those stars that do not destroy themselves in the process become one of three types of compact objects—white dwarfs, neutron stars, or black holes. Almost all of the energy available has been squeezed out of compact objects, and you find them in their final, high-density states.

You studied white dwarfs in the previous chapter. Here you will learn about the most extreme of the compact objects, and you need to compare evidence and theory with great care. Theory predicts the existence of these objects; but, by their nature, they are difficult to detect. To confirm the theories, astronomers searched for real objects that can be identified as having the properties predicted by theory. That is, they looked for real neutron stars and real black holes. It was a difficult quest but ultimately successful.

16-1 Neutron Stars

A **NEUTRON STAR,** containing a little over 1 solar mass compressed to a radius of about 10 km, can be left behind by a type II supernova. For comparison, a white dwarf has about the same mass but is about the size of Earth. A neutron star's density is so high that physicists calculate that this material is stable only as a fluid of neutrons. Theory predicts that such an object would spin a number of times a second, be nearly as hot at its surface as the inside of the sun, and have a magnetic field a trillion times stronger than Earth's. Two questions should occur to you immediately. First, how could any theory predict such a bizarre object? And second, do neutron stars really exist?

Theoretical Prediction of Neutron Stars

Neutrons were discovered in a laboratory in February 1932. Only two years later, in January 1934, Caltech astronomers Walter Baade and Fritz Zwicky published a seminal paper. They showed that some novae in historical records were much more luminous than the rest and suggested that these were caused by the collapse of a massive star's core resulting in an explosion they named a supernova. The core of the star, they proposed, would form a small and tremendously dense sphere of neutrons, and Zwicky coined the term *neutron star.*

Over the following years, scientists applied the principles of quantum mechanics to see if such an object was indeed possible. Neutrons spin in much the way that electrons do, which means that neutrons must obey the Pauli exclusion principle. In other words, if neutrons are packed together tightly enough, they can become degenerate just as electrons do. White dwarfs are supported by degenerate electrons, and quantum mechanics predicts that an even denser mass of neutrons could support itself by the pressure of degenerate neutrons.

How does the core of a collapsing star become a mass of neutrons? Nuclear physics provides an explanation. As the supernova begins, the core collapses inward. If the collapsing core is more massive than the Chandrasekhar limit of 1.4 solar masses, then it cannot reach stability as a white dwarf because the weight is too great to be supported by degenerate electrons. The collapse of the core continues, and the atomic nuclei are broken apart by gamma rays. Almost instantly, the increasing density forces the freed protons to combine with electrons and become neutrons by this reaction:

$$e + p \rightarrow n + \nu$$

A result of the production of each neutron is a neutrino. You learned in the previous chapter that the burst of neutrinos (ν) during a supernova explosion helps blast the envelope of the star away. The star's core is left behind as a neutron star.

Which stars produce neutron stars as remnants? As you saw in the previous chapter, a star of 8 solar masses or less can lose enough mass to die by forming a planetary nebula and then leaving behind a white dwarf. More massive stars also lose mass rapidly, but model calculations indicate they cannot shed mass fast enough to reduce their mass below the Chandrasekhar limit, so it seems likely that they must die in supernova explosions. Theoretical calculations suggest that stars that begin life on the main sequence with roughly 8 to 20 solar masses will leave behind neutron stars. Stars even more massive than that are thought to form black holes.

How massive can a neutron star be? That is a critical question and a difficult one to answer because scientists don't know the strength of pure neutron material. They can't make such matter in the laboratory, so its properties must be predicted theoretically. The most widely accepted calculations suggest that a neutron star cannot be more massive than about 3 solar masses. If a neutron star were more massive than that, the degenerate neutrons would not be able to support the weight, and the object would collapse, presumably becoming a black hole). Two of the most massive neutron stars observed have masses of 1.94 and 2.74 solar masses, which confirms that hypothesis.

How big are neutron stars? Mathematical models predict that a neutron star should be only 10 or so kilometers in radius (■ Figure 16-1), which, combined with a typical mass,

■ **Figure 16-1**

A tennis ball and a road map illustrate the relative size of a neutron star. Such an object, containing slightly more than the mass of the sun, would fit with room to spare inside the beltway around Washington, D.C.

means it must have a density of almost 10^{15} g/cm^3. On Earth, a sugar-cube-sized lump of this material would weigh 100 million tons. This is roughly the density of an atomic nucleus, so you can think of a neutron star as matter with absolutely all of the empty space squeezed out of it.

Simple physics, the same physics you have used in previous chapters to understand normal stars, predicts that neutron stars should be hot, spin rapidly, and have strong magnetic fields. You have seen that contraction heats the gas in a star. As gas particles fall inward, they pick up speed, and when they collide, their high speeds become thermal energy. The sudden collapse of the core of a massive star to a radius of 10 km should heat it to a trillion degrees. Astronomers calculate that neutron stars should initially cool rapidly from that astonishingly high temperature because neutrinos can escape from their entire volume and carry energy away. After a few years neutron stars will cool down to a mere million degrees or so, and neutrinos won't be produced in bulk. From that point on, neutron stars should cool slowly because the heat can escape only from the surface, and neutron stars are so small they have little surface from which to radiate.

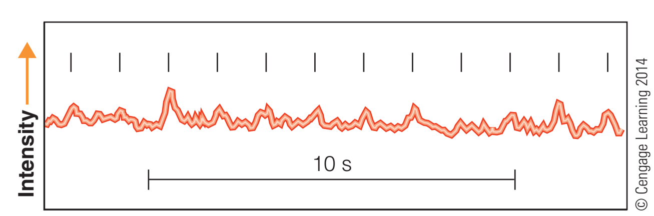
The principle of conservation of angular momentum predicts that neutron stars should spin rapidly. All stars rotate because they form from swirling clouds of interstellar matter. As a star collapses, it must rotate faster because it conserves angular momentum. Recall the example of an ice skater spinning slowly with her arms extended and then speeding up as she pulls her arms closer to her body (look back to Figure 15-11). In the same way, a collapsing star must spin faster as it pulls its matter closer to its axis of rotation. If the sun collapsed to a radius of 10 km, its rate of rotation would increase from once every 25 days to over 20 times a second. You might expect the collapsed core of a massive star to rotate in the range of 10 to 100 times a second.

It isn't hard to understand how a neutron star could have a powerful magnetic field. All stars have magnetic fields, and some have fields 1000 times stronger than the sun's. The field is frozen into the ionized gas of the star, and when it collapses into a neutron star, the field is squeezed down, concentrated, and made as much as a billion times stronger. Neutrons have no charge, so they can't anchor a magnetic field, but they are unstable and spontaneously decay into protons and electrons. Those particles are immediately forced by the pressure to merge back into neutrons, but at any moment about 10 percent of the particles in a neutron star are protons and electrons, which have electrical charges, so powerful electric currents can flow in a neutron star and sustain a very strong magnetic field.

Theory predicts the properties of neutron stars, but it also predicts that they should be difficult to observe. Neutron stars are very hot, so from your understanding of blackbody radiation you can predict they will radiate most of their energy in the gamma-ray and X-ray part of the spectrum, radiation that could not be observed in the 1940s and 1950s because astronomers could not yet put telescopes above Earth's atmosphere. Also, the small surface areas of neutron stars mean that they will be faint objects. Consequently, astronomers of the mid-20th century were not surprised that none of the newly predicted neutron stars had been found. Neutron stars were, at that point, entirely theoretical objects.

The Discovery of Pulsars

In November 1967, Jocelyn Bell, a graduate student at Cambridge University in England, found a peculiar pattern in the data from a radio telescope. Unlike other radio signals from celestial bodies, this was a series of regular pulses (■ Figure 16-2). At first she and the leader of the project, Anthony Hewish, thought the signal



■ **Figure 16-2**

The 1967 detection of regularly spaced pulses in the output of a radio telescope led to the discovery of pulsars. This record of the radio signal from the first pulsar, CP 1919, contains regularly spaced pulses (marked by ticks). The period is 1.33730119 seconds.

was interference from earthly sources, but they found it day after day in the same place in the sky. Clearly, it was celestial in origin.

Another possibility, that it came from a distant civilization, led them to consider naming it LGM, for Little Green Men. But within a few weeks, the team found three more objects in other parts of the sky pulsing with different periods. The objects were clearly natural, and the team dropped the name LGM in favor of **pulsar**—a contraction of *pulsing star.* The pulsing radio source Bell had observed with her radio telescope was the first known pulsar.

As more pulsars were found, astronomers argued over their nature. The periods ranged from 0.033 to 3.75 seconds, and each one was nearly as exact as an atomic clock. However, months of observation showed that many of the periods were slowly growing longer by a few billionths of a second per day. Whatever produced the regular pulses had to be highly precise, nearly as exact as an atomic clock, but also gradually slowing down.

It was easy to eliminate possibilities. Pulsars could not be ordinary stars. A normal star, even a small white dwarf, is much too big to pulse that fast. Nor could a star with a hot spot on its surface spin fast enough to produce the pulses. Even a small white dwarf would fly apart if it spun 30 times a second.

The pulses themselves gave the astronomers another clue. The pulses last only about 0.001 second, placing an upper limit on the size of the object producing the pulse. If a white dwarf blinked on and then off in that interval, you would not see a 0.001-second pulse. That's because the point on the white dwarf closest to Earth would be about 6000 km closer to you, and light from that spot would arrive 0.022 second before the light from the bulk of the white dwarf. As a result its short blink would be smeared out into a longer pulse. This is an important principle in astronomy—an object cannot change its brightness appreciably in an interval shorter than the time light takes to cross its diameter. If pulses from pulsars are no longer than 0.001 second, then the objects cannot be larger than 300 km (190 miles) in diameter, and could be much smaller.

Only a neutron star is small enough to be a pulsar. In fact, a neutron star is so small that it couldn't pulsate slowly enough, but it can spin as fast as 1000 times a second without flying apart. The missing link between pulsars and neutron stars was found in 1968, when astronomers discovered a pulsar at the heart of the Crab Nebula (look back to Figure 15-16). The Crab Nebula is a supernova remnant, and theory predicts that some supernovae leave behind neutron stars. The short pulses and the discovery of the pulsar in the Crab Nebula were strong evidence that pulsars are neutron stars.

A Model Pulsar

As you have noticed before, scientists often work by building a model of a natural phenomenon—not a physical model made of plastic and glue, but an intellectual conception of how nature works in a specific instance. The astronomer's model may be limited and incomplete, but it helps them organize their understudying.

The modern model of a pulsar has been called the **lighthouse model** and is shown in **The Lighthouse Model of a Pulsar** on pages 348–349. Notice three important points:

1. A pulsar does not pulse but rather emits beams of radiation that sweep around the sky as the neutron star rotates. If the beams do not sweep over Earth, the pulses will not be detectable by Earth's radio telescopes.
2. The mechanism that produces the beams involves extremely high energies and is not fully understood.
3. Modern space telescopes observing from above Earth's atmosphere can image details around young neutron stars and even locate isolated neutron stars whose beams of electromagnetic radiation do not sweep over Earth.

Neutron stars are complicated objects with extreme conditions, and modern astronomers need to use both general relativity and quantum mechanics to try to understand them. Nevertheless, astronomers know enough to tell the life story of pulsars.

The Evolution of Pulsars

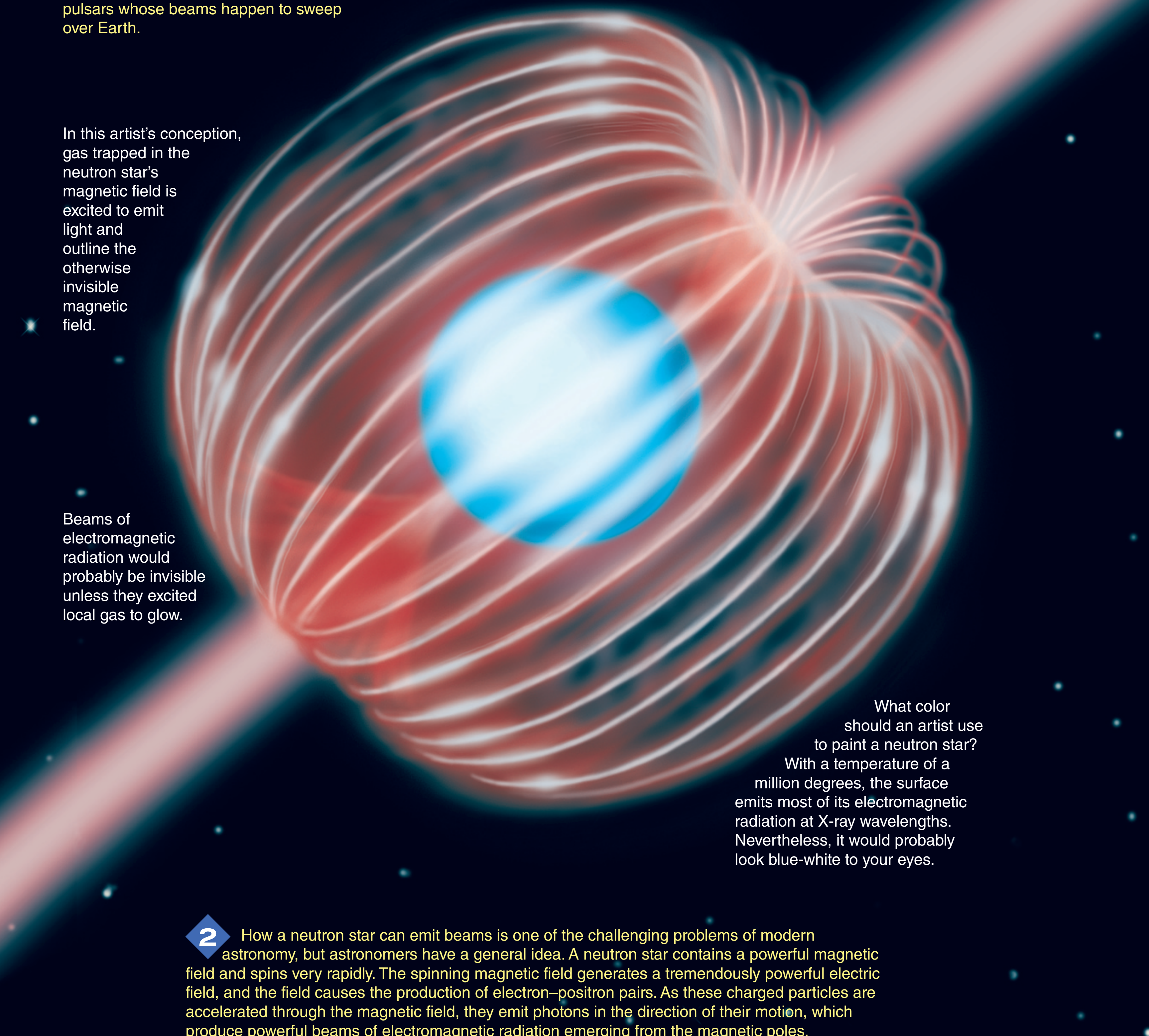
When a pulsar first forms, it is spinning fast, perhaps a hundred times a second. The energy it radiates into space comes from its energy of rotation, so as it blasts beams of radiation outward, its rotation slows. The average pulsar is apparently only a few million years old, and the oldest are about 10 million years old. Presumably, older neutron stars rotate too slowly to generate detectable radio beams.

You can expect that a young neutron star should emit powerful beams of radiation. The Crab Nebula is an example. Only about 950 years old, the Crab pulsar is so powerful it emits photons all across the electromagnetic spectrum, at radio, infrared, visible, X-ray, and gamma-ray wavelengths (■ Figure 16-3). Careful measurements of its brightness with high-speed instruments show that it blinks twice for every rotation. When one beam sweeps almost directly over Earth, astronomers detect a strong pulse. Half a rotation later, the edge of the other beam brushes over Earth, and astronomers detect a weaker pulse.

Only the most energetic pulsars produce short-wavelength photons and thus pulse at visible wavelengths. The Crab Nebula pulsar is young and powerful, and it produces visible-light pulses. So does another young pulsar called the Vela pulsar (located in the Southern Hemisphere constellation Vela). Compared with most pulsars, the Vela pulsar is fast, pulsing about 11 times a second and, like the Crab Nebula pulsar, is located inside a supernova remnant. Its age is estimated at a relatively young 20,000 to 30,000 years.

The Lighthouse Model of a Pulsar

1 Astronomers think of pulsars not as pulsing objects but rather as objects emitting beams. As they spin, the beams sweep around the sky; when a beam sweeps over Earth, observers detect a pulse of radiation. Understanding the details of this lighthouse model is a challenge, but the implications are clear. Although a neutron star is only a few kilometers in radius, it can produce powerful beams. Also, observers tend to notice only those pulsars whose beams happen to sweep over Earth.



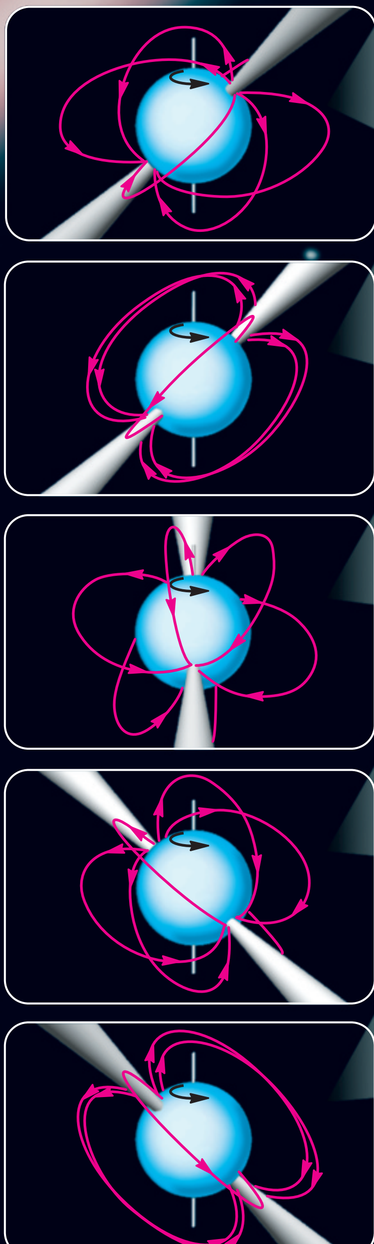
In this artist's conception, gas trapped in the neutron star's magnetic field is excited to emit light and outline the otherwise invisible magnetic field.

Beams of electromagnetic radiation would probably be invisible unless they excited local gas to glow.

What color should an artist use to paint a neutron star? With a temperature of a million degrees, the surface emits most of its electromagnetic radiation at X-ray wavelengths. Nevertheless, it would probably look blue-white to your eyes.

2 How a neutron star can emit beams is one of the challenging problems of modern astronomy, but astronomers have a general idea. A neutron star contains a powerful magnetic field and spins very rapidly. The spinning magnetic field generates a tremendously powerful electric field, and the field causes the production of electron–positron pairs. As these charged particles are accelerated through the magnetic field, they emit photons in the direction of their motion, which produce powerful beams of electromagnetic radiation emerging from the magnetic poles.

Neutron Star Rotation with Beams



As in the case of Earth, the magnetic axis of a neutron star can be inclined to its rotational axis.

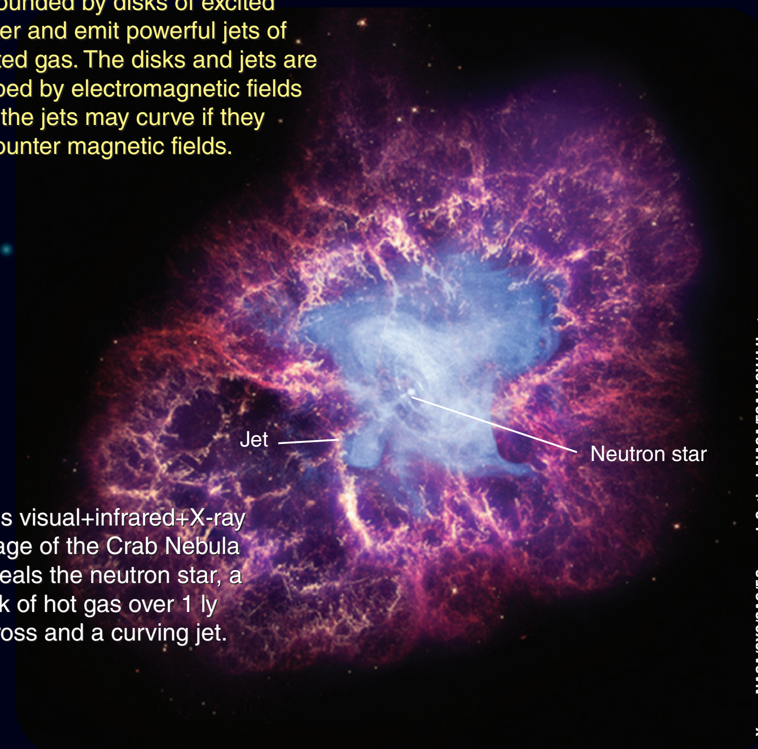
The rotation of the neutron star will sweep its beams around like beams from a lighthouse.

While a beam points roughly toward Earth, observers detect a pulse.

While neither beam is pointed toward Earth, observers detect no energy.

Beams may not be as exactly symmetric as in this model.

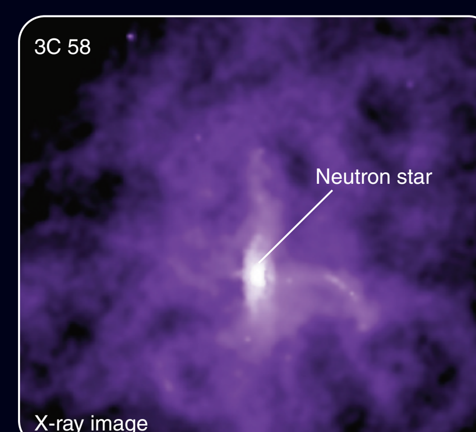
3 X-ray observations of young pulsars show that they are surrounded by disks of excited matter and emit powerful jets of excited gas. The disks and jets are shaped by electromagnetic fields and the jets may curve if they encounter magnetic fields.



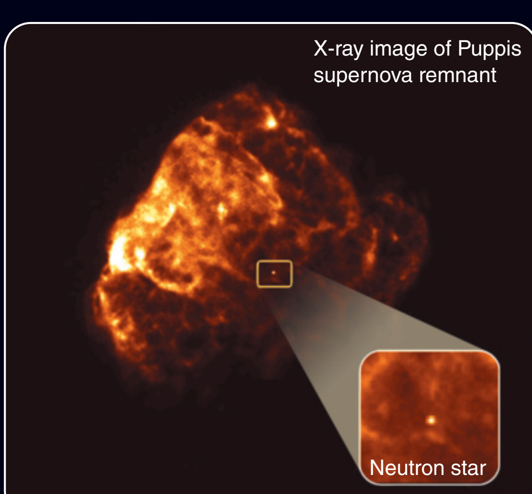
This visual+infrared+X-ray image of the Crab Nebula reveals the neutron star, a disk of hot gas over 1 ly across and a curving jet.

X-ray: NASA/CXC/SAO/F.Seward; Optical: NASA/ESA/ASU/J.Hester & A.Loll; Infrared: NASA/JPL-Caltech/Univ. Minn./R.Gehrz

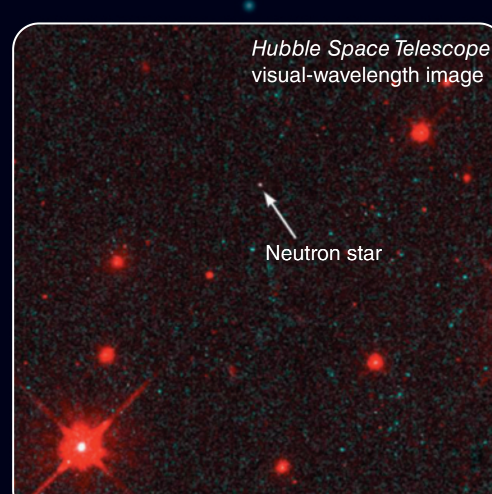
Pulsar 3C 58 at right was produced by the supernova seen in the year 1181. It pulses 15 times per second, is surrounded by a disk, and is ejecting jets in both directions



NASA/CXC/SAO/P.Slane et al.



NASA



Fred Walter (State University of New York at Stony Brook) and NASA

3a If a pulsar's beams do not sweep over Earth, observers detect no pulses, and the neutron star is difficult to find. A few such objects are known, however. The Puppis A supernova remnant is about 4000 years old and contains a point source of X-rays thought to be a neutron star. The isolated neutron star in the right-hand image has a temperature of 700,000 K.

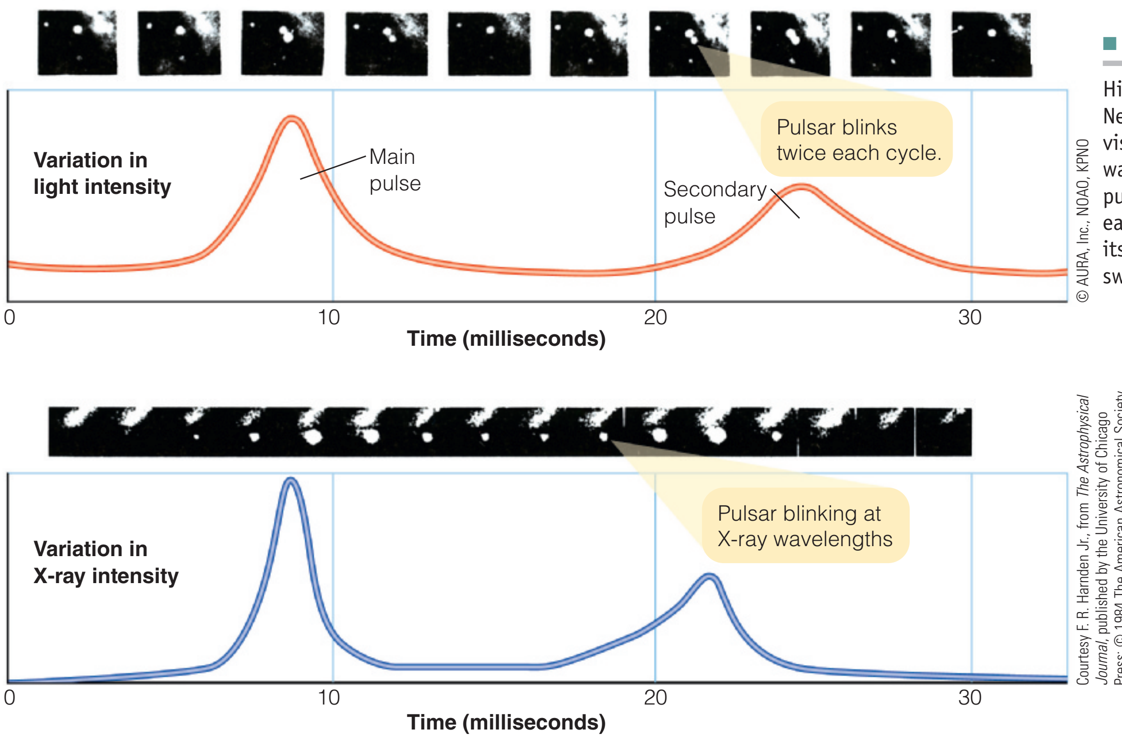


Figure 16-3

High-speed images of the Crab Nebula pulsar show it pulsing at visual wavelengths and at X-ray wavelengths. The period of pulsation is 33 milliseconds, and each cycle includes two pulses as its two beams of unequal intensity sweep over Earth.

The *Fermi Gamma-Ray Space Telescope* has detected a new kind of pulsar that pulses in pure gamma rays. It is not clear how these pulses are produced, but because gamma rays are extremely short-wavelength, high-energy photons, the process must involve very high energies. At least one of those gamma-ray pulsars is located inside a supernova remnant that is only about 10,000 years old.

The electromagnetic energy in the beams is just a small part of the energy emitted by a pulsar. Roughly 99.9 percent of the energy flowing away from a pulsar is carried as a **pulsar wind** of high-speed atomic particles. This can produce small, high-energy nebulae near a young pulsar (■ Figure 16-4).

You might expect to find all pulsars inside supernova remnants, but the statistics must be examined with care. Many supernova remnants probably do contain pulsars, but their beams never sweep over Earth. Also, some pulsars move through space at high velocity (■ Figure 16-5), quickly leaving their supernova remnants behind. Evidently supernova explosions can occur asymmetrically, perhaps because of the violent turbulence in the exploding core, and that can kick a neutron star away with a high velocity through space. Some supernovae probably occur in binary systems and fling the two stars apart at high velocity. In any case, pulsars are known to have such high velocities that many probably escape the disk of our galaxy. Finally, it seems that pulsars remain detectable for 10 million years or so, but a supernova remnant cannot survive more than about 50,000 years before it is mixed into the interstellar medium. For all these reasons, you should not be surprised that most pulsars are not in supernova remnants and that most supernova remnants do not contain pulsars.

Astronomers conclude that the 1987 explosion of Supernova 1987A formed a neutron star because a burst of neutrinos was detected passing through Earth a few hours before the visible explosion was first detected. Theory predicts that the collapse of a massive star's core into a neutron star produces such a burst of neutrinos, so the detection of the neutrinos is evidence that the supernova produced a neutron star. At first the neutron star is expected to be hidden at the center of the expanding shells of gas ejected into space, but as the gas expands and thins, astronomers eventually should be able to detect the neutron star. Even if its beams don't sweep over Earth, astronomers should eventually be able to detect its X-ray and gamma-ray emission. Although no neutron star has yet been detected, astronomers continue to watch that site, hoping to see a newborn pulsar.

One reason pulsars are so fascinating is the extreme conditions found in spinning neutron stars. To see natural processes of even greater violence, you have only to look at pulsars in binary systems.

Binary Pulsars

Nearly 2000 pulsars are now known, and some are located in binary systems. These pulsars are of special interest because astronomers can learn more about the neutron star by studying the orbital motions of the binary. Also, in some cases, mass can flow from the companion star onto the neutron star, and that produces high-energy violence.

The first binary pulsar was discovered in 1974 when astronomers Joseph Taylor and Russell Hulse noticed that the pulse period

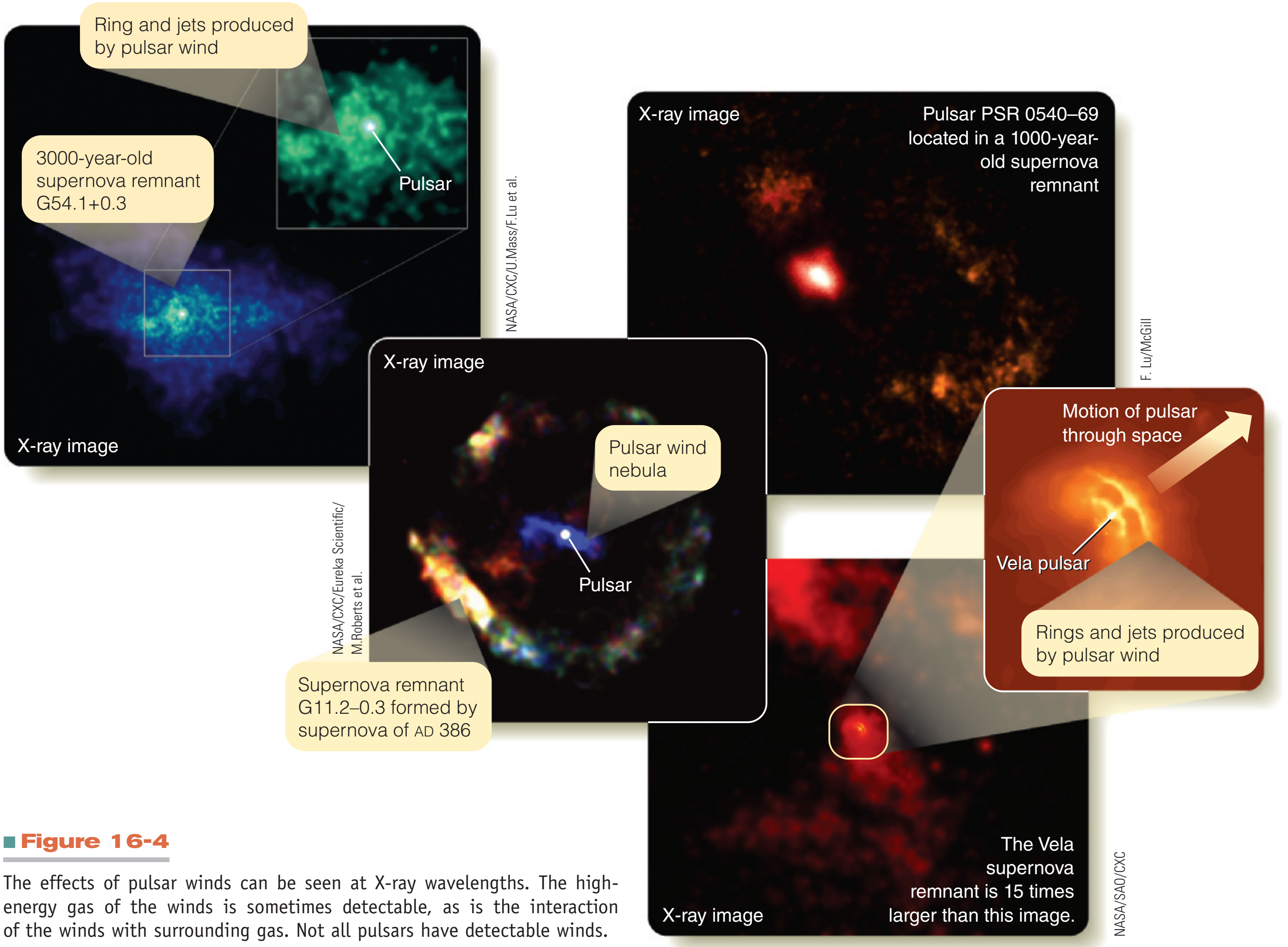


Figure 16-4

The effects of pulsar winds can be seen at X-ray wavelengths. The high-energy gas of the winds is sometimes detectable, as is the interaction of the winds with surrounding gas. Not all pulsars have detectable winds.

of the pulsar PSR 1913+16 was changing. The period first grew longer and then grew shorter in a cycle that took 7.75 hours. Thinking of the Doppler shifts seen in spectroscopic binaries, the radio astronomers realized that the pulsar had to be in a binary system with an orbital period of 7.75 hours. When the orbital motion of the pulsar carries it away from Earth, astronomers see the pulse period lengthen slightly—a redshift. Then, when the pulsar rounds its orbit and approaches Earth, they see the pulse period shorten slightly—a blueshift. From these changing Doppler shifts, Taylor and Hulse calculated the radial velocity of the pulsar around its orbit just as if it were a spectroscopic binary star. The resulting graph of radial velocity versus time was then analyzed to find the shape of the pulsar's orbit (■ Figure 16-6). The analysis of PSR 1913+16 showed that the binary system consists of two neutron stars, one of which is "silent," separated by a distance roughly equal to the radius of our sun.

Yet another surprise was hidden in the motion of PSR 1913+16. In 1916, Einstein's general theory of relativity described

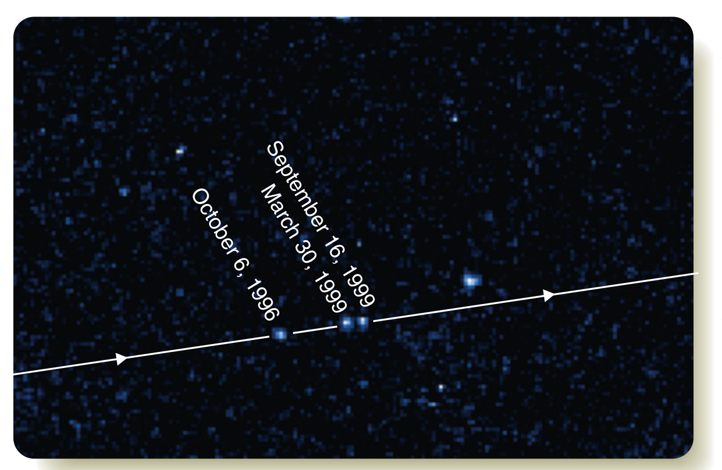


Figure 16-5

Many neutron stars have high velocities through space. Here the neutron star known as RX J185635-3754 was photographed on three different dates as it rushed past background stars.

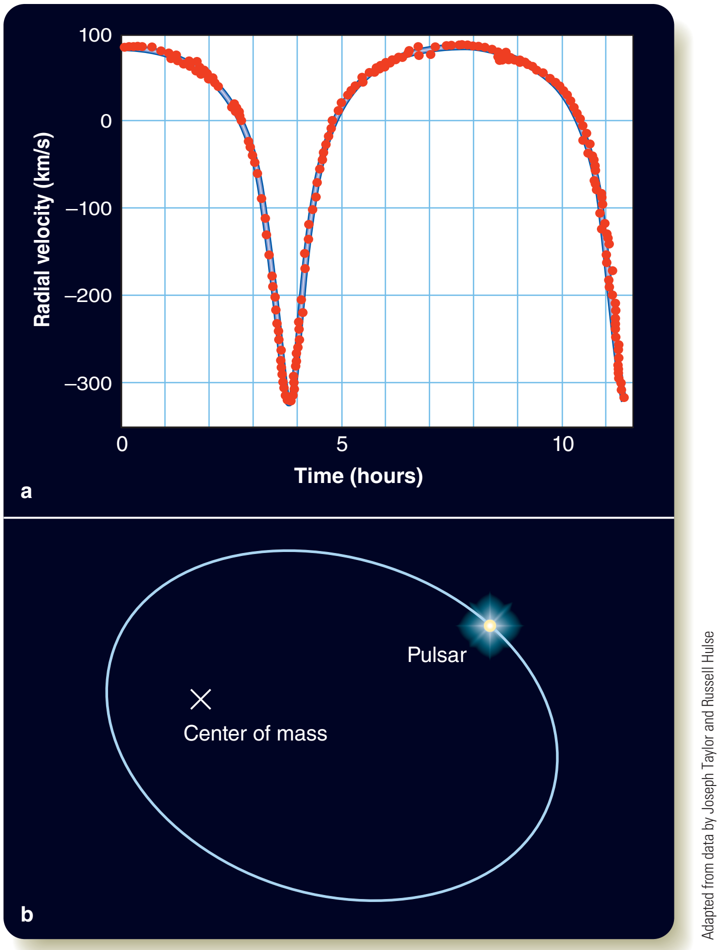


Figure 16-6

(a) The radial velocity of pulsar PSR 1913+16 can be found from the Doppler shifts in its pulsation. (b) Analysis of the radial velocity curve allows astronomers to determine the pulsar's orbit. Here the center of mass does not appear to be at a focus of the elliptical orbit because the orbit is inclined to the line of sight from Earth.

gravity as a curvature of space-time. Einstein realized that any rapid change in a gravitational field should spread outward at the speed of light as **gravitational radiation.** Gravity waves have not been detected yet, but Taylor and Hulse were able to show that the orbital period of the binary pulsar is slowly growing shorter because the stars are radiating orbital energy away as gravitational radiation and gradually spiraling toward each other. (Normal binary stars are too far apart and orbit too slowly to emit significant gravitational radiation.) Taylor and Hulse won the Nobel Prize in 1993 for their work confirming general relativity using binary pulsars.

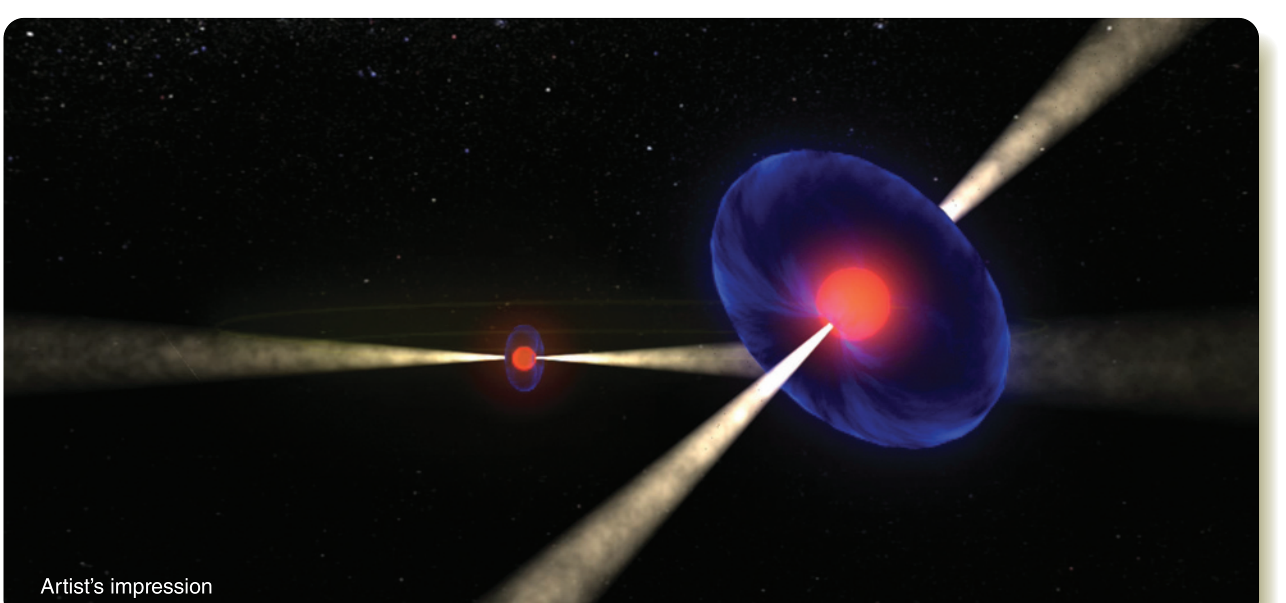
Dozens of pulsars have been found orbiting stars of various kinds; by analyzing the Doppler shifts in their pulse periods, astronomers can estimate the masses of the neutron stars. Typical masses are about 1.4 solar masses, as predicted by theory.

In 2004, astronomers discovered a binary containing two pulsars with periods of 0.022 s and 2.8 s. Because the magnetic fields eclipse each other, astronomers can study the system in detail. General relativity predicts that the objects are losing energy by gravitational radiation and will probably merge in about 85 million years, presumably triggering a violent explosion.

The gravitational fields of neutron stars are so strong they can be sites of titanic violence. An astronaut stepping onto the surface of a neutron star would be instantly smooshed into a layer only 1 atom thick. If you dropped an apple onto the surface of a neutron star from a distance of 1 AU, it would hit with an impact equivalent to a 1-megaton nuclear bomb. In general, a particle falling

Figure 16-7

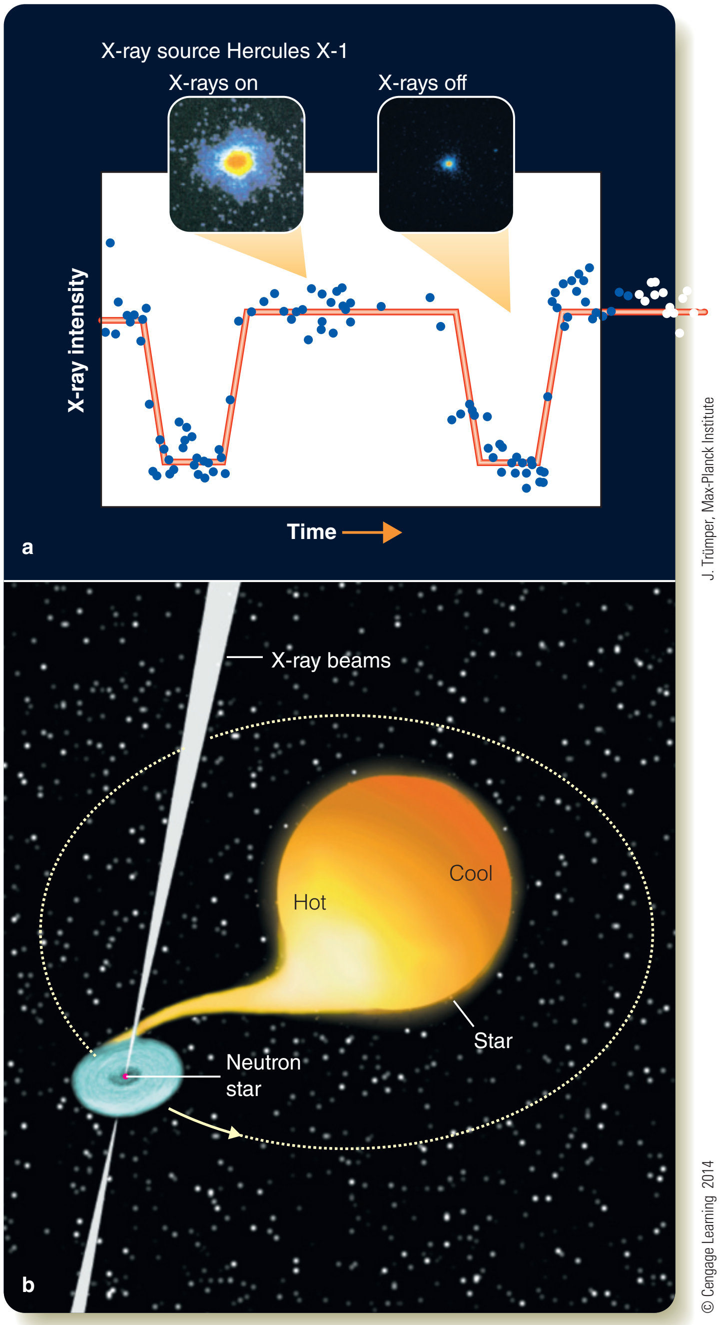
In the double pulsar, two neutron stars orbit each other in a plane that is edge on as seen from Earth. The resulting eclipses allow astronomers to study the magnetic fields, and the decay of their orbits caused by gravitational radiation may provide further tests of general relativity.



from a large distance to the surface of a neutron star will release energy equivalent to 0.2 mc^2, where m is the particle's mass at rest and c is the speed of light. Even a small amount of matter flowing from a companion star to a neutron star can generate high temperatures and release X-rays and gamma rays.

Hercules X-1 is an example of such an active system, containing a 2-solar-mass star and a neutron star that orbit each other with a period of 1.7 days (■ Figure 16-8). Matter flowing from the normal star into an accretion disk around the neutron star reaches temperatures of millions of degrees and emits a powerful X-ray glow. Interactions of the gas with the neutron star's magnetic field produce beams of X-rays that sweep around with the rotating neutron star (Figure 16-8b). Earth receives a pulse of X-rays every time a beam points this way. The X-rays shut off completely every 1.7 days when the neutron star is eclipsed behind the normal star. Hercules X-1 has many different high-energy processes going on simultaneously, but this quick sketch serves to illustrate how complex and powerful such binary systems are during mass transfer.

The X-ray source 4U 1820-30 illustrates another way neutron stars can interact with normal stars. In this system, a neutron star and a white dwarf orbit their center of mass with a period of only 11 minutes (■ Figure 16-9a and b). The separation between the two objects is only about one third the distance between Earth and the moon, smaller than a main sequence star. To explain how such a very close pairing of stellar remnants could originate from what once was an ordinary binary star system, theorists suggest that a neutron star collided with a giant star and went into an orbit inside the star. (Recall the low density of the outer envelope of giant stars.) The neutron star would gradually eat away the giant star's envelope from the inside, and the star would collapse into a white dwarf. Matter still flows from the white dwarf into an accretion disk and then down to the surface of the neutron star (Figure 16-9c), where it accumulates in a degenerate layer until it ignites helium fusion to produce a burst of X-rays. Objects called **X-ray bursters** are thought to be such binary systems involving mass transferred to a neutron star, and the bursts repeat each time a large enough layer of degenerate fuel accumulates. Notice the similarity between this mechanism and that responsible for novae (look back again to Chapter 15).

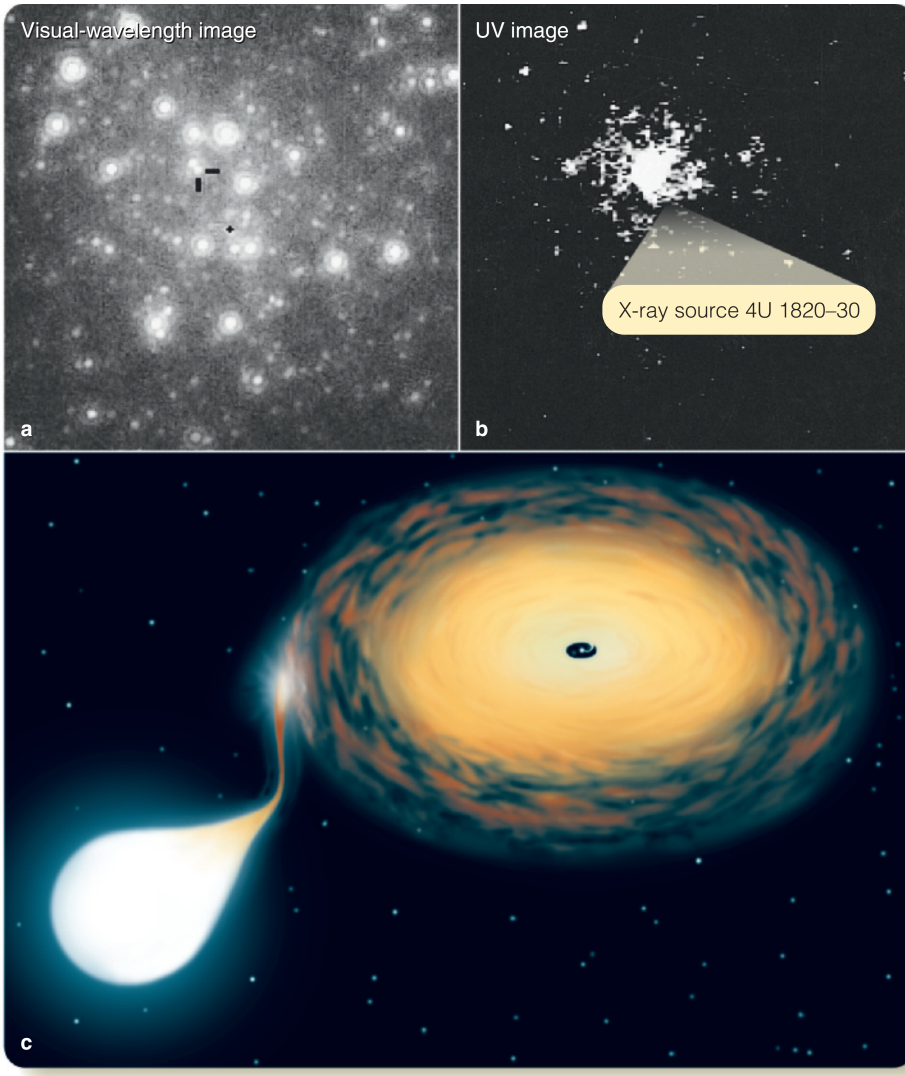


■ **Figure 16-8**

Sometimes the X-ray pulses from Hercules X-1 are on, and sometimes they are off. A graph of X-ray intensity versus time looks like the light curve of an eclipsing binary. (b) In Hercules X-1, matter flows from a star into an accretion disk around a neutron star producing X-rays, which heat the near side of the star to 20,000 K compared with only 7000 K on the far side. X-rays turn off when the neutron star is eclipsed behind the star.

The Fastest Pulsars

Your knowledge of pulsars suggests that newborn pulsars should blink rapidly, and old pulsars should blink slowly. In fact, the handful that blink the fastest may be quite old. One of the fastest known pulsars is cataloged as Terzan 5ad in the constellation Sagittarius. It pulses 716 times a second and is slowing down only slightly. The energy stored in the rotation of a neutron star at this rate is equal to the total energy of a supernova explosion, so it seemed difficult at first to explain this pulsar. It now appears that Terzan 5ad is an old neutron star that has gained mass and rotational energy from a companion in a binary system. Like water hitting a mill wheel, the matter falling on the neutron star has spun it up to 716 rotations per second. With its weak magnetic field, it slows down very gradually and will continue to spin rapidly for a very long time.



NASA/STScI

NASA/STScI

Dana Berry, STScI/NASA

© Cengage Learning 2014

Figure 16-9

(a) At visible wavelengths, the center of star cluster NGC 6624 is crowded with stars. (b) In the ultraviolet, one object stands out, an X-ray source consisting of a neutron star orbiting a white dwarf. (c) An artist's conception shows matter flowing from the white dwarf into an accretion disk around the neutron star.

A number of other very fast pulsars have been found. They are generally known as **millisecond pulsars** because their pulse periods and therefore their periods of rotation are almost as short as a millisecond (0.001 s). This rapid rotation produces some fascinating physics. If a neutron star 10 km in radius spins 716 times a second, as does Terzan 5ad, then its equator must be traveling about 45,000 km/s. That's fast enough to flatten the neutron star into an ellipsoidal shape. The fastest known pulsar, XTE J1739-285, spins 1122 times a second. It appears to have been spun up by matter flowing from its companion star and is very near the breakup speed for an object composed of pure neutrons.

"Show me," say scientists; and, in the case of neutron stars, the evidence seems so strong that astronomers have great confidence that such objects really do exist. Of course, you can never prove that a theory is absolutely true (**How Do We Know? 16-1**), but the evidence for neutron stars is so strong that astronomers have great confidence that they really do exist. Other theories that describe how they form and evolve and how they emit beams of radiation are less certain, but continuing observations at many wavelengths are expanding astronomers' understanding of these last embers of massive stars. In fact, precise observations have turned up some objects no one expected.

The hypothesis that millisecond pulsars were spun up by mass transfer from a companion star is quite reasonable, but scientists demand evidence, and evidence has been found. One example, the X-ray source XTE J1751-305, is a pulsar with a period of 2.3 milliseconds. X-ray observations show that it is gaining mass from a companion star. The orbital period is only 42 minutes, and the mass of the companion star is only 0.014 solar mass. The evidence suggests that this neutron star has devoured all but the last morsel of its binary partner.

Astronomers wondered about a few millisecond pulsars that do not have companions. Were they produced by some process other than mass transfer? Pulsar B1957+20, also known as the Black Widow (Figure 16-10) seems to show that they all might once have been members of binary systems. The Black Widow has a period of 1.6 milliseconds and is orbited by a brown dwarf companion. There is no evidence of current mass transfer. However, spectra of the system show that blasts of radiation and high-energy particles from the neutron star are now evaporating the companion. When the companion is completely gone, presumably a new solitary millisecond pulsar will be left behind.

Pulsar Planets

Because a pulsar's period is so precise, astronomers can detect tiny variations by comparing their observations with atomic clocks. When astronomers checked pulsar PSR B1257+12, they found variations in the period of pulsation much like those caused by the orbital motion of a binary pulsar (Figure 16-11a). However, in the case of PSR B1257+12, the variations were much smaller, and when they were interpreted as Doppler shifts, it became evident

How Do We Know? | 16-1

Theories and Proof

Why do astronomers say that a theory is confirmed but never say that it is proven? What scientists mean by the word *theory* is a hypothesis that has "graduated" to being confidently considered a well-tested truth. You can think of a hypothesis as equivalent to having a suspect in a criminal case, and a theory as equivalent to the trial being finished and someone being convicted of the crime.

Of course, no matter how many tests and experiments you conduct, you can never prove that any scientific theory is absolutely true. It is always possible that the next observation you make will disprove the theory. And it is unfortunately sometimes true that innocent people go to jail and guilty people are free, although sometimes, with further evidence, those legal mistakes can be fixed.

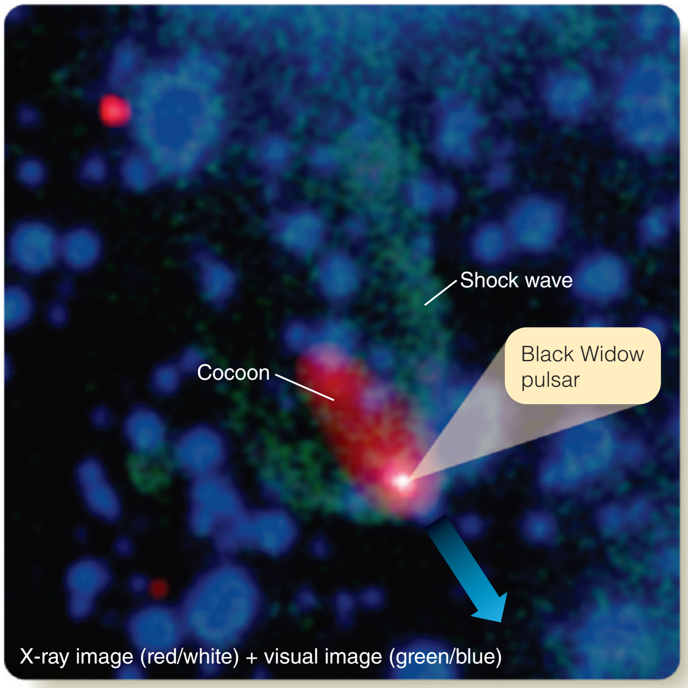
There have always been hypotheses about why the sun is hot. Only a century ago, most astronomers accepted the hypothesis that the sun was hot because gravity was making it contract. In the late 19th century, geologists showed that Earth was much older than the sun could be if it was powered by gravity, so the gravity hypothesis had to be wrong. It wasn't until 1920 that Sir Arthur Eddington suggested the sun is powered somehow by the energy in atomic nuclei. In 1938 the German-American astrophysicist Hans Bethe showed how nuclear fusion could power the sun. He won the Nobel Prize in 1967.

The fusion hypothesis is now so completely confirmed that it is fair to call it a theory. No one will ever go to the center of the sun, so you can't *prove* the fusion theory is right. Many observations and model calculations support this theory, and in Chapter 7 you learned about further evidence provided by the neutrinos that have been detected coming from the sun's core. Nevertheless, there remains some tiny possibility that all the observations and models are misunderstood and that the theory will be overturned by some future discovery.

There is a great difference between a theory in the colloquial sense of a far-fetched guess and a scientific theory that has undergone decades of testing and confirmation with observations, experiments, and models. But no theory can ever be proven absolutely true. It is up to you as a consumer of knowledge and a responsible citizen to distinguish between a flimsy guess and a well-tested theory that deserves to be treated like truth—at least pending further information.

SOHO/MDI

Technically it is still a theory, but astronomers have great confidence that the sun gets its power from nuclear fusion.



© Cengage Learning 2014; X-ray image: NASA/CXC/ASTRON/B. Stappers et al.; Optical image: AAO/J. Bland-Hawthorn & H. Jones

Figure 16-10

The Black Widow pulsar and its companion star are moving rapidly through space, creating a shock wave like the bow wave of a speedboat. The shock wave confines high-energy particles shed by the pulsar into an elongated cocoon (red).

that the pulsar was being orbited by at least two objects with planetlike masses of 4.1 and 3.8 Earth masses. The gravitational tugs of these planets make the pulsar wobble about the center of mass of the system by no more than 800 km, producing the observed tiny changes in period (Figure 16-11b).

Astronomers greeted this discovery with both enthusiasm and skepticism. As usual, they looked for ways to test the hypothesis. Simple gravitational theory predicts that planets in the same system should interact and slightly modify each other's orbits. When the data were analyzed, that interaction was found,

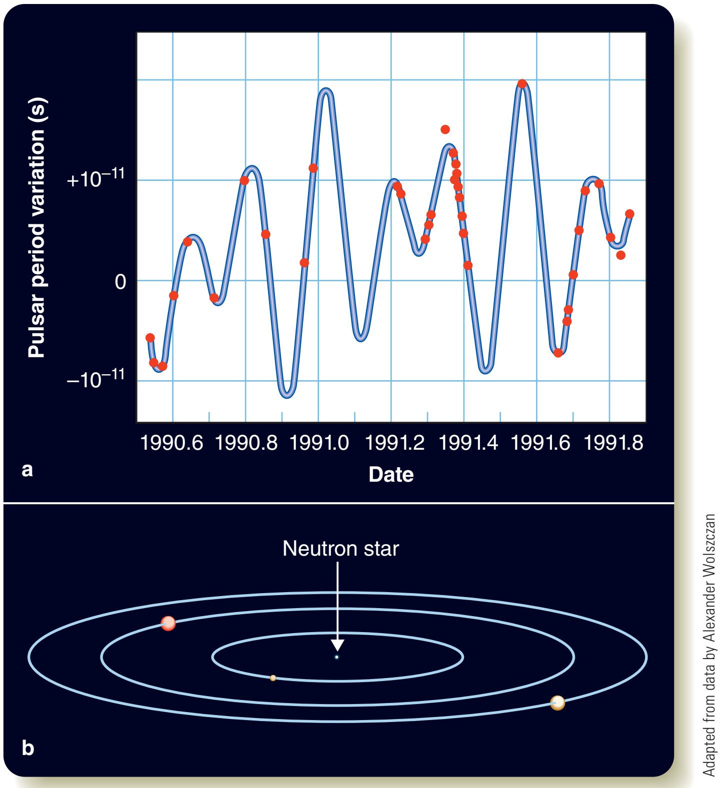


Figure 16-11

(a) The dots in this graph are observations showing that the period of pulsar PSR 1257+12 varies from its average value by a fraction of a billionth of a second. The blue line shows the variation that would be produced by planets orbiting the pulsar. (b) As the planets orbit the pulsar, they cause it to wobble by less than 800 km, a distance that is invisibly small in this diagram.

further confirming the existence of the planets. In fact, later data revealed the presence of a third planet with only 1/40 the mass of Earth, about twice the mass of Earth's moon. This illustrates the astonishing precision of studies based on pulsar timing.

Astronomers wonder how a neutron star can have planets. The three planets that orbit PSR B1257+12 are closer to the pulsar than Venus is to the sun. Any planets that orbited a star that closely would have been absorbed or vaporized when the star expanded to become a supergiant. Furthermore, the supernova explosion would have suddenly reduced the mass of the star and allowed any orbiting planets to escape from their orbits. So how can planets exist there? One suggestion is that these planets are the remains of a stellar companion that was devoured by the neutron star. In fact, PSR B1257+12 spins very fast (161 pulses per second), suggesting that it was spun up in a binary system. However, the Spitzer Space Telescope observing in the infrared has detected a ring of gas and dust around a different rapidly spinning neutron star. If supernova explosions can leave such rings of material behind, then perhaps planets can form from the accumulation of matter in the rings.

PSR B1257+12 is not unique. Another planet has been found orbiting a pulsar that is part of a binary system with a white dwarf in a very old star cluster. The characteristics of this system indicate, however, that the planet may have been captured rather than being debris from the supernova explosion that made the neutron star. Planets probably orbit other neutron stars, and small shifts in the timing of the pulses may eventually reveal their presence.

You can imagine what these worlds might be like. Formed from the remains of elderly stars, they might have chemical compositions richer in heavy elements than Earth. You can imagine visiting these worlds, landing on their surfaces, and hiking across their valleys and mountains. Above you, the neutron star would glitter in the sky, a tiny point of light.

SCIENTIFIC ARGUMENT

Why are neutron stars easier to detect at X-ray wavelengths?
This argument draws together a number of ideas you know from previous chapters. First, recall that a neutron star is very hot because of the heat released when it contracts to a radius of 10 km. It could easily have a surface temperature of 1,000,000 K, and Wien's law (look back to Chapter 6) tells you that such an object will radiate most intensely at a very short wavelength—X-rays and gamma rays. Normal stars are much cooler and emit only weak X-rays unless they have hot accretion disks. At visual wavelengths, stars are bright, and neutron stars are faint, but at X-ray wavelengths, the neutron stars stand out from the crowd.

Now build a new argument as if you were seeking funds for a research project. **What observations would you make to determine whether a newly discovered pulsar was young or old, single or a member of a binary system, alone or accompanied by planets?**

16-2 Black Holes

You have now studied white dwarfs and neutron stars, two of the three end states of dying stars. Now it's time to think about the third end state—black holes.

Although the physics of black holes is difficult to discuss without using sophisticated mathematics, simple logic is sufficient to predict that they should exist. The problem is to confirm that they are real. What objects observed in the heavens could be real black holes? More difficult than the search for neutron stars, the quest for black holes has nevertheless met with success.

You can begin by considering a simple question. How fast must an object travel to escape from the surface of a celestial body?

Escape Velocity

Suppose you threw a baseball straight up. How fast must you throw it if it is not to come down? Of course, gravity will always pull back on the ball, slowing it, but if the ball is traveling fast enough to start with, it will never come to a stop and fall back. Such a ball will escape from Earth.

In Chapter 4 you learned that the escape velocity is the initial velocity an object needs to escape from a celestial body (■ Figure 16-12). Whether you are discussing a baseball leaving Earth or a particle escaping a collapsing star, escape velocity depends on two things, the mass of the celestial body and the distance from the center of mass to the escaping object. If the celestial body has a large mass, its gravity is strong, and you need a high velocity to escape, but if you begin your journey farther from the center of mass, the velocity needed is less. For example, to escape from Earth, a spaceship would have to leave Earth's surface at 11 km/s (25,000 mph), but if you could launch spaceships from the top of a tower 1000 miles high, the escape velocity would be only 10 km/s (22,000 mph).


© Cengage Learning 2014

■ **Figure 16-12**

Escape velocity, the velocity needed to escape from a celestial body, depends on mass. The escape velocity at the surface of a very small body would be so low you could jump into space. Earth's escape velocity is much larger, about 11 km/s (25,000 mph).

If you could make an object massive enough or small enough, its escape velocity could be greater than the speed of light. Relativity says that nothing can travel faster than the speed of light, so even photons, which have no mass, would be unable to escape. Such a small, massive object could never be seen because light could not leave it.

Long before Einstein and relativity, the Reverend John Mitchell, a British gentleman astronomer, realized this particular consequence of Newton's laws of gravity and motion. In 1783, he pointed out that an object 500 times the radius of the sun but of the same density would have an escape velocity greater than the speed of light. Then, "all light emitted from such a body would be made to return towards it." Mitchell didn't know it, but he was talking about a black hole.

Schwarzschild Black Holes

If the core of a star contains more than 3 solar masses when it collapses, no force can stop it. It cannot stop collapsing when it reaches the density of a white dwarf because degenerate electrons cannot support that weight, and it cannot stop when it reaches the density of a neutron star because not even degenerate neutrons can support that weight. No force remains to stop the object from collapsing to zero radius.

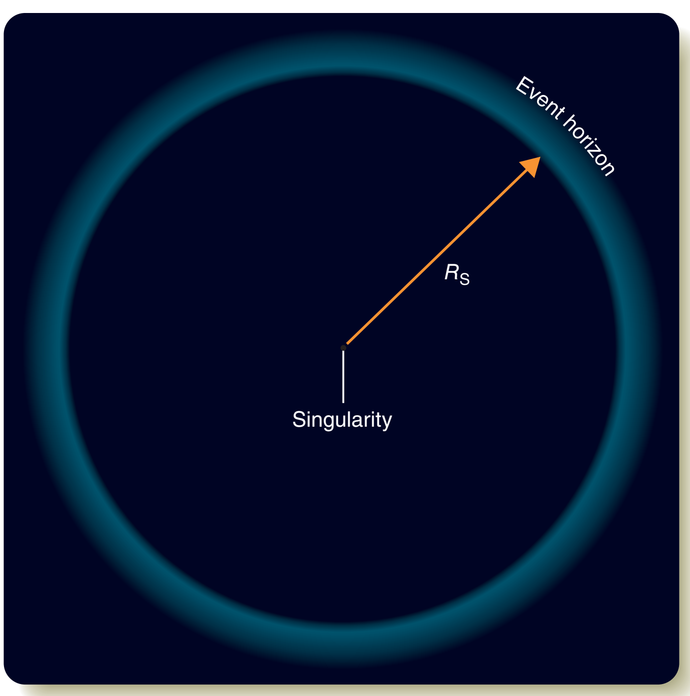
As an object collapses, its density and the strength of its surface gravity increase. If an object collapses to zero radius, its density and gravity become infinite. Mathematicians call such a point a **singularity,** but in physical terms it is difficult to imagine an object of zero radius. Some theorists believe that a singularity is impossible and that the laws of quantum physics must somehow halt the collapse at some subatomic radius roughly 10^{20} times smaller than a proton. Astronomically, it seems to make little difference.

If the contracting core of a star becomes small enough, the escape velocity in the region of space around it is so large that no light can escape. This means you can receive no information about the object or about the region of space near it. Because it emits no light, such a region is called a **black hole.** If the core of an exploding star collapsed into a black hole, the expanding outer layers of the star could produce a supernova remnant, but the core would vanish without a trace.

To understand black holes, you should consider relativity. In 1916, Albert Einstein published a mathematical theory of space and time that became known as the general theory of relativity. Einstein treated space and time as a single entity called space-time. His equations showed that gravity could be

described as a curvature of space-time, and almost immediately the astronomer Karl Schwarzschild found a way to solve Einstein's equations to describe the gravitational field around a single, nonrotating, electrically neutral lump of matter. That solution contained the first general relativistic description of a black hole, and nonrotating, electrically neutral black holes are now known as Schwarzschild black holes. In recent decades, theorists such as Roy Kerr and Stephen Hawking have found ways to apply the sophisticated mathematical equations of the general theory of relativity and quantum mechanics to describe charged, rotating black holes. For this discussion, the differences are minor, and you may proceed as if all black holes were Schwarzschild black holes.

Schwarzschild's solution shows that if matter is packed into a small enough volume, then space-time curves back on itself. Objects can follow paths that lead into the black hole, but no path leads out, so nothing can escape. Because not even light can escape, the inside of the black hole is totally beyond the view of an outside observer. The **event horizon** is the boundary between the isolated volume of space-time and the rest of the universe, and the radius of the event horizon is called the **Schwarzschild radius, R_S**. A collapsing stellar core must shrink inside its Schwarzschild radius to become a black hole (■ Figure 16-13).



■ Figure 16-13

A black hole forms when an object collapses to a small size (perhaps to a singularity) and the escape velocity becomes so great light cannot escape. The boundary of the black hole is called the event horizon because any event that occurs inside is invisible to outside observers. The radius of the black hole R_S is the Schwarzschild radius.

Although Schwarzschild's work was highly mathematical, his conclusion is quite simple. The Schwarzschild radius depends only on the mass of the object:

$$R_S = \frac{2GM}{c^2}$$

In this simple formula, G is the gravitational constant, M is the mass (in kilograms), and c is the speed of light (in meters per second). A bit of arithmetic shows that a 1-solar-mass black hole has a Schwarzschild radius of 3 km, a 10-solar-mass black hole has a Schwarzschild radius of 30 km, and so on.

Every object has a Schwarzschild radius determined by its mass, but not every object is a black hole. For example, Earth has a Schwarzschild radius of about 1 cm, meaning that it could become a black hole if you squeezed it smaller than that radius. Fortunately, Earth will not collapse spontaneously to become a black hole because the strength of the rock and metal in its interior is sufficient to support its weight. Only extinguished stellar cores more massive than about 3 solar masses can form black holes under the sole influence of their own gravity.

It is a **Common Misconception** to think of black holes as giant vacuum cleaners that will eventually suck in everything in the universe. A black hole is just a gravitational field, and at a reasonably large distance its gravity is no greater than that of a normal object of similar mass. If the sun were replaced by a 1-solar-mass black hole, the orbits of the planets would not change at all. ■ Figure 16-14 illustrates this by representing gravitational fields as curvature of the fabric of space-time. Normal uncurved space-time is represented by a flat plane, and the presence of a mass such as a star curves the plane to produce a depression. The extreme curvature around a black hole produces a deep funnel-shaped surface in this graphic representation. You can see from the graphs that the gravity of a black hole becomes extreme only when you approach close to it.

Now you can check off another **Common Misconception** that may strike you as silly. Because of special effects in movies and TV, some people think black holes should actually look like funnels. Of course, the graphs of the strength of gravity around black holes look like funnels, but black holes themselves are not shaped like funnels. If you could approach a black hole, you might be able to see hot gas swirling inward, but you wouldn't be able to see the black hole itself.

Leaping into a Black Hole

Before you can search for real black holes, you need to understand what theory predicts about a black hole. To explore that idea, you can imagine leaping, feet first, into a Schwarzschild black hole.

If you were to leap into a black hole of a few solar masses from a distance of 1 AU, the gravitational pull would not be very large, and you would fall slowly at first. Of course, the longer you fell and the closer you came to the center, the faster you would

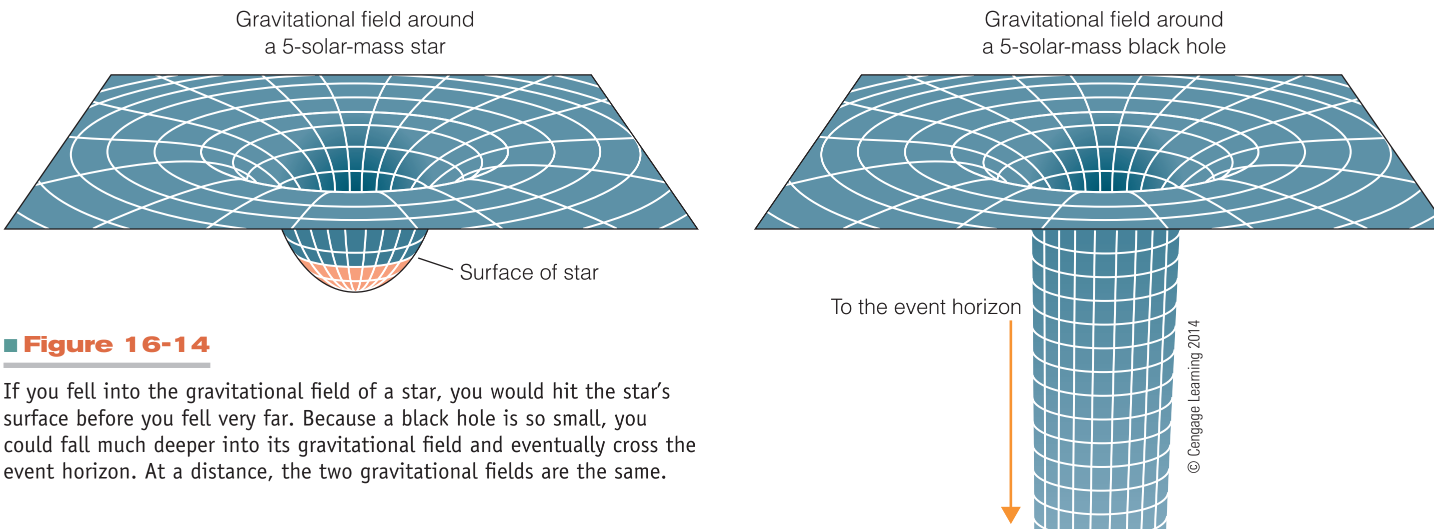


Figure 16-14

If you fell into the gravitational field of a star, you would hit the star's surface before you fell very far. Because a black hole is so small, you could fall much deeper into its gravitational field and eventually cross the event horizon. At a distance, the two gravitational fields are the same.

travel. Your wristwatch would tell you that you fell for about two months by the time you reached the event horizon.

Your friends who stayed behind would see something different. They would see you falling more and more slowly as you came closer to the event horizon because, as explained by general relativity, time slows down in curved space-time. This is known as **time dilation.** In fact, your friends would never actually see you cross the event horizon. To them you would fall more and more slowly until you seemed hardly to move. Generations later, your descendants could focus their telescopes on you and see you still inching closer to the event horizon. You, however, would have sensed no slowdown and would conclude that you had reached the event horizon after about two months.

Another relativistic effect would make it difficult to see you with normal telescopes. As light travels out of a gravitational field, it loses energy, and its wavelength grows longer. This is known as the **gravitational redshift.** Although you would notice no effect as you fell toward the black hole, your friends would need to observe at longer and longer wavelengths to detect you.

While these relativistic effects seem merely peculiar, other effects would be quite unpleasant. If you were falling feet first, you would feel your feet, which would be closer to the black hole, being pulled in more strongly than your head. This is a tidal force, and at first it would be minor. But as you got closer to the black hole, the tidal force would become very large. Another tidal force would compress you as both your left and your right side fell toward the center of the black hole. For any black hole with a mass like that of a star, the tidal forces would crush you sideways and stretch you lengthwise long before you reached the event horizon (Figure 16-15). The friction from

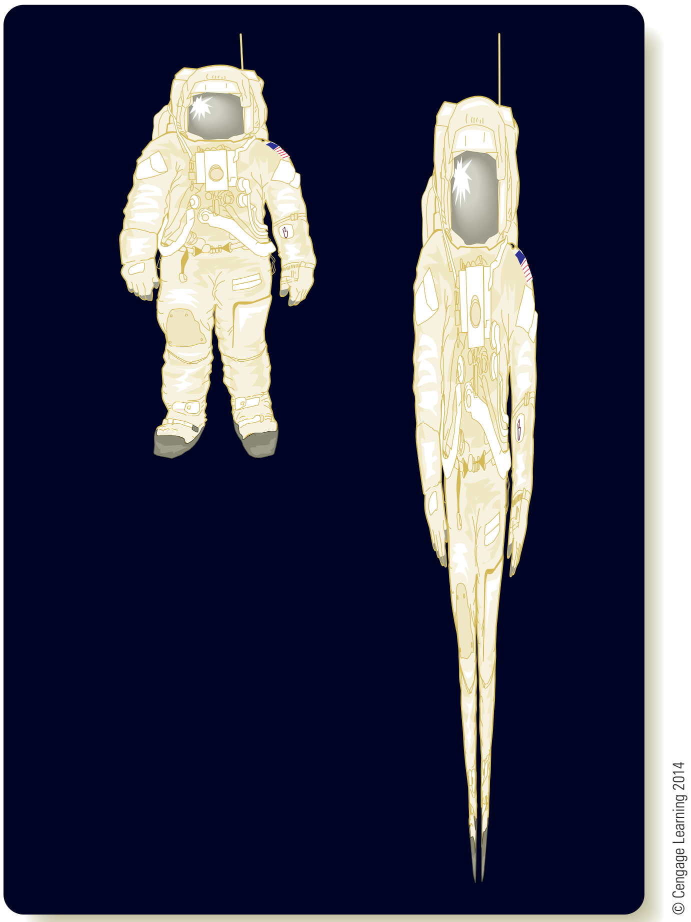


Figure 16-15

Leaping feet-first into a black hole. A person of normal proportions (left) would be distorted by tidal forces (right) long before reaching the event horizon around a typical black hole of stellar mass. Tidal forces would stretch the body lengthwise while compressing it laterally. Friction from this distortion would heat the body to high temperatures.

How Do We Know? 16-2

Checks on Fraud in Science

How do you know scientists aren't just making stuff up? The unwritten rules of science make fraud difficult, and the way scientists publish their research makes it almost impossible. Scientists depend on each other to be honest, but they also double-check everything.

For example, all across North America, black-capped chickadees sing the same quick song. Some people say it sounds like *Chick-a-dee-dee-dee,* but others say it sounds like *Hey-sweetie-sweetie-sweetie*. You could invent tables of data and publish a paper reporting that you had recorded chickadees around Ash Lake in northern Minnesota that sing a backward song: *Sweetie-sweetie-sweetie-hey*. Experts on brain development and animal learning would be amazed, and your research might secure you praise from your colleagues, a job offer at a prestigious university, or a generous grant—but only if you could get away with it.

The first step in your scheme would be to publish your results in a scientific journal. Because the journal's reputation rests on the accuracy of the papers it publishes, the editor sends all submitted papers to one or more experts for peer review. Those world experts on chickadees would almost certainly notice things wrong with your made-up data tables. On their recommendation, the editor would probably refuse to publish your paper.

Even if your faked data fooled the peer reviewers, you would probably be found out once the paper was published. Experts on bird song would read your paper and flock to Ash Lake to study the bird songs themselves. By the next spring, you would be found out—and the journal would be forced to publish an embarrassing retraction of your article.

One of the rules of science is that good results must be repeatable. Scientists routinely repeat the work of others, not only to check the results but as a way to start a new research topic. When someone calls a news conference and announces a new discovery, other scientists begin asking, "How does this fit with other observations? Has this been checked? Has this been peer-reviewed?" Until a result has been published in a peer-reviewed journal, scientists treat it with extra care.

Fraud isn't unheard of in science. But because of peer review and the requirement of repeatability in science, bad research, whether the result of carelessness or fraud, is usually exposed quickly.

Steve and Dave Maslowski/Photo Researchers, Inc.

Chickadees always sing the same song. Hey-Sweetie-Sweetie-Sweetie.

such severe distortions of your body would heat you to millions of degrees, and you would emit X-rays and gamma rays. (Needless to say, this would render you inoperative as a thoughtful observer.)

Some people have suggested that it is possible to travel through the universe by jumping into a black hole in one place and popping out of another somewhere far across space. That might make for good science fiction, but tidal forces would make it an unpopular form of transportation even if it worked. You would certainly lose your luggage.

Your imaginary leap into a black hole is not frivolous. You now know how to find a black hole: Look for a strong source of X-rays. It may be a black hole into which matter is falling and being heated.

The Search for Black Holes

Do black holes really exist? The first X-ray telescopes reached orbit in the 1970s, and that allowed astronomers to begin searching for evidence of black holes. They tried to find one or more objects that were obviously black holes. That very difficult search is a good illustration of how the unwritten rules of science help scientists understand nature (**How Do We Know? 16-2**).

A black hole alone is totally invisible because nothing can escape from its event horizon, but if matter flows into a black hole, it will whirl through an accretion disk and become hot enough to emit X-rays before it reaches the event horizon. An isolated black hole in space will not have much matter flowing into it, but a black hole in a binary system might receive a steady flow of matter transferred from the companion star. This suggests you can search for black holes by looking closely at X-ray binaries.

Some X-ray binaries such as Hercules X-1 contain a neutron star, and they will emit X-rays much as would a binary containing a black hole. You can tell the difference in two ways. If the compact object emits pulses, you know it is a neutron star because the neutron star has a solid body within which powerful magnetic fields are generated. As the neutron star rotates, the magnetic field emits beams that sweep around the sky producing pulses as they pass across Earth. A black hole inside its event horizon cannot emit regular pulses like those from a neutron star. Another clue depends on the mass of the object. If the mass of the compact object is greater than about 3 solar masses, it cannot be a neutron star; it must be a black hole.

The first X-ray binary suspected of harboring a black hole was Cygnus X-1, the first X-ray object discovered in the

artwork © Don Dixon, cosmographica.com

■ **Figure 16-16**

The X-ray source Cygnus X-1 consists of a supergiant B0 star and a compact object orbiting each other. Gas from the B0 star's stellar wind flows into the hot accretion disk around the compact object, and the X-rays astronomers detect come from the disk.

constellation Cygnus. It contains a supergiant B0 star and a compact object orbiting each other with a period of 5.6 days. Astronomers suspected that the X-rays were emitted by matter from the star flowing into the compact object. That object is invisible, but Doppler shifts in the spectrum reveal the motion of the B0 star around the center of mass of the binary. From the geometry of the orbit, astronomers were able to calculate that the mass of the compact object had to be greater than 3.8 solar masses, well above the maximum for a neutron star.

To confirm that black holes exist, astronomers needed a conclusive example, an object that couldn't be anything else. Cygnus X-1 didn't quite pass that test when it was first discovered. Perhaps the B0 star was not a normal star, and its portion of the system mass was incorrectly estimated; or perhaps the system contained a third star. Either possibility would distort the analysis. At the time, astronomers could not conclusively show that Cygnus X-1 contained a black hole.

It took years of work to understand Cygnus X-1. Further observations and analysis show that the B0 star has a mass of about 25 solar masses, and the compact object is about 10 times the mass of the sun. Astronomers conclude that matter flows from the B0 star as a strong stellar wind, and much of that matter gets caught in a hot accretion disk about five times larger in diameter than the orbit of Earth's moon. The inner few hundred kilometers of the disk have a temperature of about 2 million Kelvin–hot enough to radiate X-rays (■ Figure 16-16). The evidence is now strong that Cyg X-1 contains a black hole.

As X-ray telescopes have found many more X-ray-emitting objects, the list of black hole candidates has grown to dozens. A few of these objects are shown in ■ Table 16-1. Each candidate is a compact object surrounded by a hot accretion disk in a close X-ray binary system. Some of the binary systems are easier to analyze than others, but in the end, it has become clear that at least a few of these compact objects, including Cygnus X-1, are clearly too massive to be neutron stars and must be black holes. The evidence is now overwhelming: Black holes really do exist.

Another way to confirm that black holes are real is to search for evidence of their distinguishing characteristic—event horizons—and that search also has been successful. In one study, astronomers selected 12 X-ray binary systems, six of which seemed to contain neutron stars and six of which were thought to contain

■ **Table 16-1** | **Seven Black Hole Binaries**

Object	Star	Orbital Period	Mass of Black Hole
Cygnus X-1	B0I	5.6 days	10 $M_\odot$
LMC X-3	B3V	1.7 days	>8 $M_\odot$
A0620-00	KV	7.75 hours	11 ± 1.9 $M_\odot$
V404 Cygni	G-KV	6.47 days	12 ± 3 $M_\odot$
GRO J1655-40	F5V	2.61 days	6.9 ± 1 $M_\odot$
QZ Vul	KV	8 hours	10 ± 4 $M_\odot$
4U 1543-47	AV	1.123 days	2.7–7.5 $M_\odot$

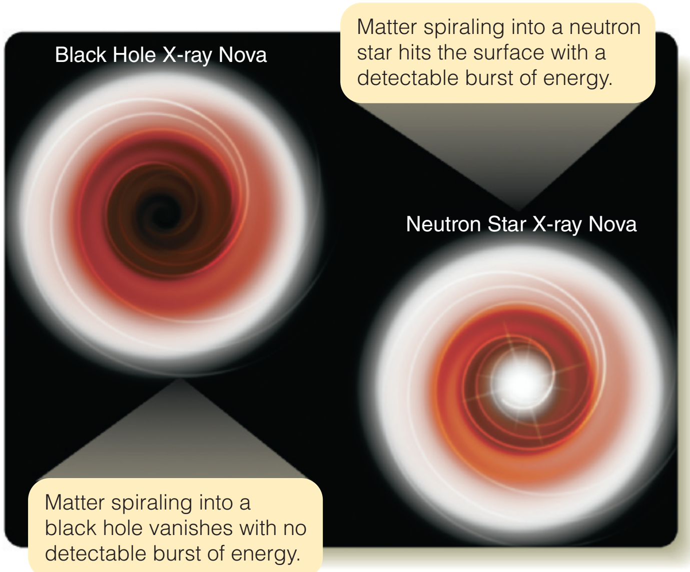


Figure 16-17

Gas spiraling into an accretion disk grows hot, and as it nears the central object, a strong gravitational redshift makes it appear redder and dimmer. Systems containing a neutron star emit bursts of energy when the gas hits the surface of the neutron star, but such bursts are not seen for systems containing black holes. In those systems, the matter vanishes as it approaches the event horizon. This is direct observational evidence of an event horizon around black holes.

black holes. Using X-ray telescopes, the astronomers monitored the systems, watching for telltale flares of energy as blobs of matter fell into the accretion disks and spiraled inward. In the six systems thought to contain neutron stars, the astronomers could detect final bursts of energy when the blobs of matter finally impacted the surfaces of the neutron stars. In the six systems suspected of containing black holes, however, the blobs of matter spiraled inward through the accretion disks and vanished without final bursts of energy. Evidently, those blobs of matter became undetectable as they approached the event horizons (■ Figure 16-17). This is dramatic evidence that event horizons are real.

The evidence shows that black holes really do exist. The problem now is to understand how these objects interact with the matter flowing into them through accretion disks to produce high-energy jets and outbursts.

SCIENTIFIC ARGUMENT

If relativistic effects slow time and prevent you from seeing matter cross the event horizon, how can infalling matter disappear without a trace?

This argument brings together observations and theory. Astronomers observed flares when matter hit the surfaces of neutron stars, but observed no flares when matter fell into a black hole. Although time slows near the event horizon, remember the gravitational red shift. Hot matter flowing into a black hole can emit X-rays, but as the matter nears the event horizon, the gravitational red shift lengthens the wavelengths dramatically. The matter vanishes, not because you see it cross the event horizon but because its photons are shifted to undetectably long wavelengths.

Now build a new argument to review a basic principle. **Why does matter become hot as it falls into a black hole?**

16-3 Compact Objects with Disks and Jets

Matter flowing onto a neutron star or onto a black hole forms an accretion disk, and that can produce some surprising phenomena. Astronomers are just beginning to understand these peculiar effects.

Jets of Energy from Compact Objects

Observations show that some compact objects are emitting jets of gas and radiation in opposite directions. These jets are similar to the bipolar outflows ejected by protostars but much more powerful. You have seen in the X-ray images on page 349 that some young pulsars, including the Crab Nebula pulsar, are ejecting jets of highly excited gas. The Vela pulsar does the same (Figure 16-4). Systems containing black holes can also eject jets. The black hole candidate GRO J1655-40 has been observed at radio wavelengths sporadically ejecting oppositely directed jets at 92 percent the speed of light.

One of the most powerful examples of this process is an X-ray binary called SS 433. Its optical spectrum shows sets of spectral lines that are Doppler shifted by about one fourth the speed of light, with one set shifted to the red and one set shifted to the blue. Furthermore, the two sets of lines shift back and forth across each other with a period of 164 days. Astronomers recognized the combination of red and blue Doppler shifts as evidence of oppositely directed jets.

Apparently, SS 433 is a binary system in which a compact object (probably a black hole) pulls matter from its companion star and forms an extremely hot accretion disk. Jets of high-temperature gas blast away from the disk in beams aimed in opposite directions. As the disk precesses, it sweeps these beams around the sky once every 164 days, and telescopes on Earth detect light from gas carried outward in both beams. One beam produces a redshift, and the other produces a blueshift.

It's not clear how an accretion disk can produce jets. Accretion disks around neutron stars and black holes are very small, spin very fast, and grow very hot. Somehow the hot gas in the disk can emit powerful beams of gas and radiation along the disk's axis of rotation. You can recognize the geometry of SS 433 in the cover illustration for this chapter (page 344). The exact process isn't well understood, but it seems to involve magnetic fields that get caught in the accretion disk and are twisted into tightly wound tubes that squirt gas and radiation out of the disk and confine it in narrow beams. Such pairs of jets are a prototype that illustrates how the gravitational field around a compact object can produce powerful beams of radiation and matter. You will meet this phenomenon again in a later chapter when you study active galaxies.

Gamma-Ray Bursts

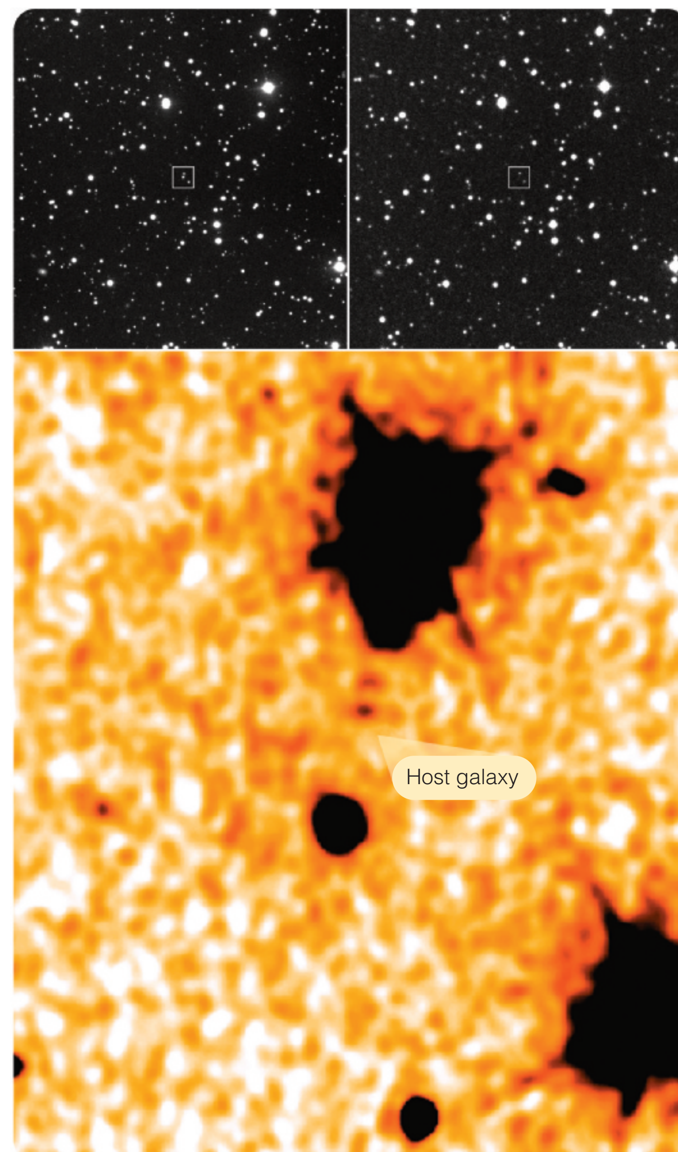
The Cold War played a minor part in the story of neutron stars and black holes. In 1963, a nuclear test ban treaty was signed, and by 1968, the United States was able to put a series of satellites in orbit to watch for nuclear tests that were violations of the treaty. A nuclear detonation emits gamma rays, so the satellites were designed to watch for bursts of gamma rays coming from Earth. The experts were startled when the satellites detected about one gamma-ray burst a day coming from space. When those data were finally declassified in 1973, astronomers realized that the bursts might be coming from neutron stars and black holes. These bursts are now known as **gamma-ray bursts** (abbreviated **GRBs**).

The Compton Gamma Ray Observatory (CGRO) reached orbit in 1991 and immediately began detecting gamma-ray bursts at the rate of a few a day. Its observations showed that the intensity of the gamma rays rises to a maximum in seconds and then fades away quickly; a burst is usually over in a few seconds to a minute.

Data from the CGRO also showed that the gamma-ray bursts were coming from all over the sky and not from any particular region. This helped astronomers eliminate some hypotheses. For example, there was a hypothesis that the gamma-ray bursts were being produced by relatively common events involving the stars in our galaxy, but these results eliminated that possibility. If the gamma-ray bursts were produced among stars in our galaxy, you would expect to see them most often along the Milky Way where there are lots of stars. That the bursts occurred all over the sky meant that they were being produced by rare events in distant galaxies.

Gamma-ray bursts are hard to study because they occur without warning and fade so quickly, but starting in 1997, new satellites were put into orbit to detect gamma-ray bursts. Their data show that there are two kinds of gamma-ray bursts. Short bursts last less than 2 seconds, but longer bursts can go on for many seconds. Specialized space observatories now can detect bursts, quickly determine their location in the sky, and immediately alert astronomers on the ground. When telescopes on Earth swiveled to image the locations of the bursts, they detected fading glows that resembled supernovae (■ Figure 16-18), suggesting that long gamma-ray bursts are produced by a certain kind of supernova explosion.

Stellar interior models indicate that a star more massive than some upper limit of about 20 solar masses can exhaust its



NASA/JPL-Caltech/P. Garnavich [Notre Dame]

ESO and NASA

© Cengage Learning 2014

■ **Figure 16-18**

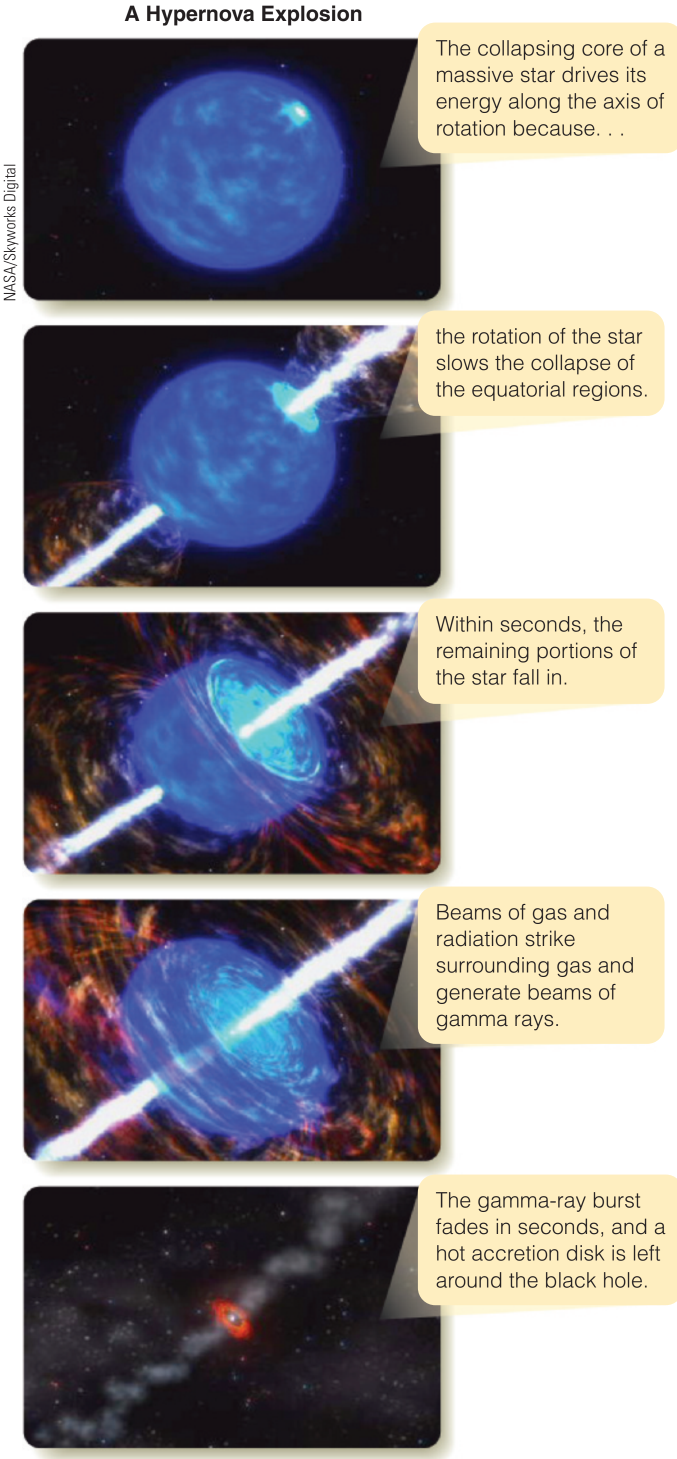
Alerted by gamma-ray detectors on satellites, observers used one of the VLT 8.2-meter telescopes on a mountaintop in Chile to image the location of a gamma-ray burst only hours after the burst. The image at top left shows the fading glow of the eruption. The image at top right, recorded 13 years before, reveals no trace of an object at the location of the gamma-ray burst. The *Hubble Space Telescope* image at bottom was recorded a year later and reveals a very faint, distant galaxy at the location of the gamma-ray burst.

nuclear fuel and collapse directly into a black hole. Models show that the collapsing star would conserve angular momentum and spin very rapidly, slowing the collapse of the equatorial parts of the star. The poles of the star would fall in quickly, and that would focus beams of intense radiation and ejected gas that would blast out along the axis of rotation. Such an eruption has been called a **hypernova** (■ Figure 16-19). If one of those beams were pointed at Earth, it could produce a powerful gamma-ray burst. Evidently the long gamma-ray bursts are produced by hypernovae.

Short gamma-ray bursts, on the other hand, don't seem to be associated with hypernovae. Some repeat, and these repeating bursts seem to be produced by neutron stars with magnetic fields 100 times stronger than that in a normal neutron star. Dubbed **magnetars,** these objects can produce bursts of gamma rays when shifts in the magnetic field break the crust of the neutron stars—causing "starquakes"—and release large amounts of energy (■ Figure 16-20). The *Fermi Gamma-Ray Space Telescope,* launched in 2008, has detected a neutron star that has eruptions as often as 100 times in 20 minutes. Another magnetar produced a burst of gamma rays that reached Earth in 1998 and was strong enough to increase the ionization of Earth's upper atmosphere noticeably, disrupting radio communication worldwide.

Not all short gamma-ray bursts are produced by magnetars. Some bursts have occurred in parts of distant galaxies where you would not expect to find the young, massive stars that produce magnetars or hypernovae, and the afterglows don't resemble fading supernovae. These bursts may be produced by the merger of two neutron stars or a neutron star and a black hole that orbited each other, radiated orbital energy as gravitational radiation, and spiraled into each other. Such a collision would cause a violent explosion as the two objects merged to form a new, or larger, black hole. The gamma-ray burst and fading afterglow from a neutron star plus black hole merger should be different from that produced by the merger of two neutron stars. Astronomers are now working to distinguish between these two kinds of short gamma-ray bursts.

In 2008 the Swift orbiting telescope detected an intense gamma-ray burst that originated in a galaxy 7.5 billion light years from Earth. The burst was so powerful that for about one minute, its visual-wavelength component was bright enough to see with the unaided eye. If you had been looking directly at it, you would have seen it appear as a star slightly brighter than those in the Little Dipper. Another powerful gamma-ray burst detected by *Swift* in 2008 originated in a galaxy 12.2 billion light-years away. In spite of the distance, it was one of the brightest gamma-ray bursts ever detected. Astronomers suspect that these bursts were produced by hypernova collapses of massive stars in which one of the jets was aimed directly at Earth.



© Cengage Learning 2014; images: NASA/Skyworks Digital

■ **Figure 16-19**

The collapse of the cores of extremely massive stars can produce hypernova explosions, which are thought to be the source of gamma-ray bursts longer than 2 seconds.

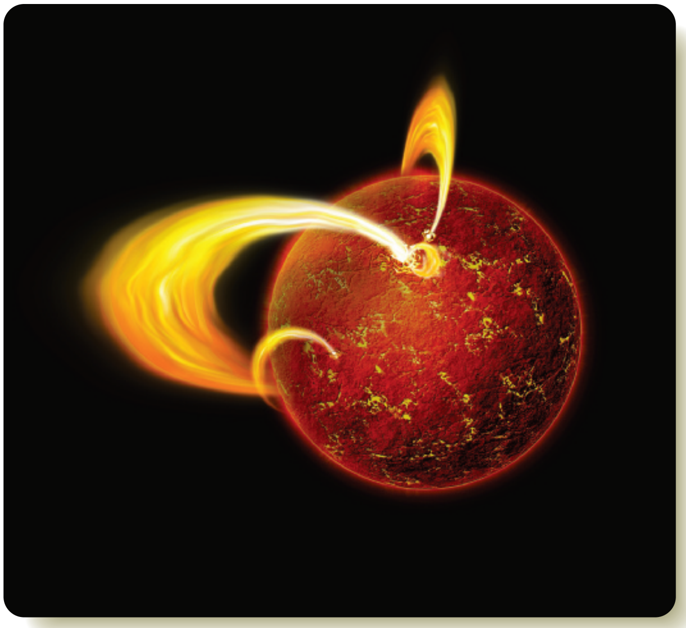

NASA/CXC/M. Weiss

Figure 16-20

Some neutron stars appear to have magnetic fields up to 1000 times stronger than those in a normal neutron star. These magnetars can produce bursts of gamma rays when shifts in the magnetic field rupture the rigid crust of the neutron star.

Could a gamma-ray burst occur near Earth? The nearest known binary pulsar is only about 2000 ly from Earth. If a gamma-ray burst occurred at that distance, the gamma rays would shower Earth with radiation equivalent to a 10,000-megaton nuclear blast, comparable to a full-scale nuclear war between superpowers. (The largest single bombs ever made released less than a hundred megatons of energy.) The gamma rays could create enough nitric oxide in the atmosphere to produce intense acid rain and also would destroy the ozone layer, exposing life on Earth to deadly levels of solar ultraviolet radiation. Gamma-ray bursts can occur relatively near the Earth as often as every few hundred million years and could be one of the causes of the mass extinctions that show up in the fossil record.

Does it surprise you that such astonishing events as merging neutron stars and hypernovae produce something so common that gamma-ray telescopes observe one or more every day? Remember that these events are so powerful they can be detected over very great distances. There may be 30,000 neutron star binaries in each galaxy, and there are billions of galaxies within range of gamma-ray telescopes. Earthlings are treated to the entire observable universe's display of these cosmic catastrophes.

What Are We? Abnormal

Look around. What do you see? A table, a chair, a tree? It's all normal stuff. The world we live in is familiar and comfortable, but astronomy reveals that "normal" isn't normal at all. The universe is, for the most part, utterly unlike anything you have ever experienced.

Throughout the universe, gravity makes clouds of gas form stars, and in turn the stars generate energy through nuclear fusion in their cores, which delays gravity's final victory. But gravity always wins. You have learned that stars of different masses die in different ways, but you have also discovered that they always reach one of three end states: white dwarfs, neutron stars, or black holes. However strange these compact objects may seem, they are common in the universe. They are normal.

The physics of compact objects is extreme and violent. You are not accustomed to objects as hot as the surface of a neutron star, and you have never experienced the environment near a black hole, where gravity is so strong it would pull you to pieces.

The universe is filled with things that are so violent and so peculiar they are almost unimaginable, but they are so common they deserve the label "normal." Next time you are out for a walk, look around and notice how beautiful Earth is and recall how unusual it is compared to the rest of the universe.

Study and Review

Summary

- When a supernova explodes, the core collapses to very small size. Theory predicts that the collapsing core cannot support itself as a white dwarf if its mass is greater than 1.4 solar masses, the Chandrasekhar limit. If its mass lies between 1.4 solar masses and about 3 solar masses, it can halt its contraction and form a **neutron star (p. 345).**
- A neutron star is supported by the pressure of the degenerate neutrons. Theory predicts that a neutron star should be about 10 km in radius, spin very fast because it conserves angular momentum as it contracts, and have a powerful magnetic field.
- **Pulsars (p. 347),** rapidly pulsing radio sources, were discovered in 1967. The **lighthouse model (p. 347)** explains pulsars as spinning neutron stars that emit beams of radiation from their magnetic poles. As they spin, they sweep the beams around the sky like lighthouses; if the beams sweep over Earth, astronomers detect pulses. The short pulses and the discovery of a pulsar in the supernova remnant called the Crab Nebula were key evidence that pulsars are neutron stars.
- A spinning neutron star slows as it radiates its energy into space. Most of the energy emitted by a pulsar is carried away as a **pulsar wind (p. 350).**
- Theory predicts that a neutron star cannot have a mass greater than about 3 solar masses. Dozens of pulsars have been found in binary systems, and those objects allow astronomers to estimate the masses of the neutron stars. Such masses are consistent with the predicted masses of neutron stars.
- Observations of the first binary containing two neutron stars revealed that the system is losing orbital energy by emitting **gravitational radiation (p. 352).**
- In some binary systems, mass flows into a hot accretion disk around the neutron star and causes the emission of X-rays. **X-ray bursters (p. 353)** are systems in which matter accumulates on the surface of the neutron star and explodes.
- The fastest pulsars, the **millisecond pulsars (p. 354),** appear to be old pulsars that have been spun up to high speed by mass flowing from binary companions.
- Planets have been found orbiting at least one neutron star. They may be the remains of a companion star that was mostly devoured by the neutron star, or they may have formed from a ring of gas and dust left orbiting the neutron star after the supernova explosion.
- If the collapsing core of a supernova has a mass greater than 3 solar masses, then it must contract to a very small size—perhaps to a **singularity (p. 357),** an object of zero radius. Near such an object, gravity is so strong that not even light can escape, and the region is called a **black hole (p. 357).**
- The outer boundary of a black hole is the **event horizon (p. 358);** no event inside is detectable. The radius of the event horizon is the **Schwarzschild radius, R_S (p. 358),** amounting to only a few kilometers for a black hole of stellar mass.
- If you were to leap into a black hole, your friends who stayed behind would see two relativistic effects. They would see your clock slow relative to their own clock because of **time dilation (p. 359).** Also, they would see your light redshifted to longer wavelengths because of the **gravitational redshift (p. 359).** You would not notice these effects, but you would feel powerful tidal forces that would deform and heat your mass until you grew hot enough to emit X-rays. Any X-rays you emitted before reaching the event horizon could escape.
- To search for black holes, astronomers must look for binary star systems in which mass flows into a compact object and emits X-rays. If the mass of the compact object is clearly greater than about 3 solar masses, then the object is presumably a black hole. A number of such objects have been located.
- Clues that an X-ray source is caused by accretion into a black hole rather than onto a neutron star are lack of sustained regular pulsations and lack of flares when blobs of material in the accretion disk arrive at the central compact object.
- Black holes and neutron stars at the center of accretion disks can eject powerful jets of gas and radiation. Such jets have been detected.
- **Gamma-ray bursts (p. 363)** appear to be related to violent events involving neutron stars or black holes. Bursts longer than 2 seconds appear to arise during **hypernovae (p. 374),** the collapse of massive stars to form black holes.
- Some short gamma-ray bursts are produced by shifts in the powerful magnetic fields in a type of pulsar called **magnetars (p. 374),** whereas other bursts probably represent the merger of binary compact objects such as neutron-star pairs or neutron-star/black-hole pairs.

Review Questions

1. How are neutron stars and white dwarfs similar? How do they differ?
2. Why is there an upper limit to the mass of neutron stars?
3. Why do you expect neutron stars to spin rapidly?
4. If neutron stars are hot, why aren't they very luminous?
5. Why do you expect neutron stars to have a powerful magnetic field?
6. Why did astronomers conclude that pulsars actually could not be pulsating stars?
7. What does the short length of pulsar pulses tell you?
8. How does the lighthouse model explain pulsars?
9. What evidence can you cite that pulsars are neutron stars?
10. Why would astronomers naturally assume that the first discovered millisecond pulsar was relatively young?
11. How can a neutron star in a binary system generate X-rays?
12. If the sun has a Schwarzschild radius, why isn't it a black hole?
13. How can a black hole emit X-rays?
14. What evidence can you cite that black holes really exist?
15. How can mass transfer into a compact object produce jets of high-speed gas? X-ray bursts? Gamma-ray bursts?
16. Discuss the possible causes of gamma-ray bursts.
17. **How Do We Know?** Why do scientists say that a hypothesis or theory is confirmed, but do not say it is proven?
18. **How Do We Know?** How does peer review make fraud rare in science?

Discussion Questions

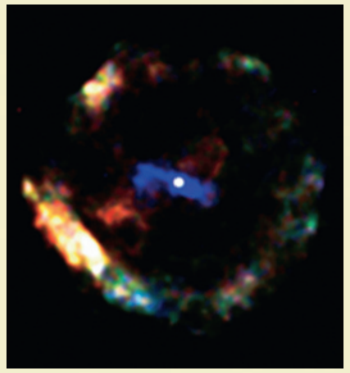
1. In your opinion, has the link between pulsars and neutron stars been sufficiently tested to be called a theory, or should it be called a hypothesis? What about the existence of black holes?
2. Why wouldn't an accretion disk orbiting a giant star get as hot as an accretion disk orbiting a compact object?

Problems

1. If a neutron star has a radius of 10 km and rotates 716 times a second, what is the speed of the surface at the neutron star's equator as a fraction of the speed of light?
2. Suppose that a neutron star has a radius of 10 km and a temperature of 1,000,000 K. How luminous is it? (*Hint:* Use the luminosity-temperature-radius relation, Chapter 13.)
3. A neutron star and a white dwarf have been found orbiting each other with a period of 11 minutes. If their masses are typical, what is their average separation? Compare the separation with the radius of the sun, 7×10^5 km. (*Hints:* Use the version of Kepler's third law for binary stars, Chapter 13; make sure you express quantities in units of AU, solar masses, and years.)
4. If the accretion disk around a neutron star has a radius of 2×10^5 km, what is the orbital velocity of a particle at its outer edge? (*Hint:* Use the formula for circular orbit velocity, Chapter 4.)
5. What is the escape velocity at the surface of a typical neutron star? (*Hints:* The formula for circular orbit velocity is in Chapter 4; escape velocity equals circular orbital velocity times the square root of 2, which is about 1.41.)
6. If Earth's moon were replaced by a typical neutron star, what would the angular diameter of the neutron star be as seen from Earth? (*Hint:* Use the small-angle formula, Chapter 3.)
7. If the inner accretion disk around a black hole has a temperature of 1,000,000 K, at what wavelength will it radiate the most energy? What part of the spectrum is this in? (*Hint:* Use Wien's law, Chapter 6.)
8. What is the orbital period of a bit of matter in an accretion disk 2×10^5 km from a 10-solar-mass black hole? (*Hint:* Use the circular orbit velocity formula, Chapter 4.)
9. If an X-ray binary consists of a 20-solar-mass star and a neutron star orbiting each other every 13.1 days, what is their average separation? (*Hints:* Use the version of Kepler's third law for binary stars, Chapter 13; make sure you express quantities in units of AU, solar masses, and years.)

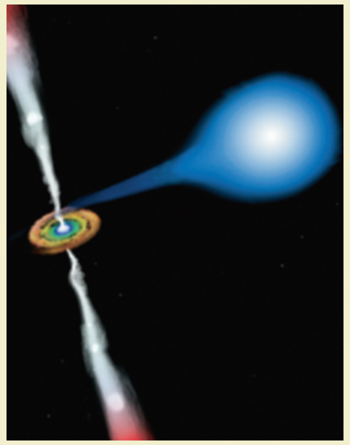
Learning to Look

1. The X-ray image at the right shows the supernova remnant G11.2–0.3 and its central pulsar in X-rays. The blue nebula near the pulsar is caused by the pulsar wind. How old do you think this system is? Discuss the appearance of this system a million years from now.



NASA/CXC/Eureka Scientific/M. Roberts et al.

2. What is happening in the artist's impression at the right? How would you distinguish between a neutron star and a black hole in such a system?



CXC/M. Weiss

Great Debates

1. *Jocelyn Bell.* Jocelyn Bell discovered pulsars. Bell helped build the radio telescope that was used to discover the pulsars, and only she operated the telescope, recorded the data, and analyzed the data. Tony Hewish, her PhD advisor, supervised. In addition to finding the original anomaly in the data, Bell found other similar pulses, confirming the anomaly was not human-made. She wrote her PhD thesis with the pulsar discovery in an appendix. However, when the research paper on the discovery of pulsars was submitted to *Nature*, Tony Hewish was listed as first author, and Bell was listed as the second of five coauthors. The Nobel Prize in Physics for the discovery of pulsars was awarded to Tony Hewish. Jocelyn Bell did not share in the Nobel Prize because awards of this caliber are not given to research students. Should the status requirement of the Nobel Prize recipient be lifted? Should Jocelyn Bell have been awarded a share in the Nobel Prize?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Use information you find in your research on those who have received the Nobel Prize and how the selection process works.
   c. Cite your sources.
2. *Pulsar Planets?* Two leading suggestions on how pulsar planets form are as follows: (1) pulsar planets form in the supernova explosion ring around the pulsar; and (2) pulsar planets are acquired as spoils when the pulsar devours the planets' host star. Which suggestion is more correct?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
3. *Gravitational Waves.* As Einstein explained, a rapid change in a gravitational field should spread outward at the speed of light. These gravity waves have not yet been detected because the waves are very weak. What useful information can mankind gain from this discovery, if the waves are detected? If the gravity wave is detected, is the information gained worth a Nobel Prize in Physics? How would you vote if you were on the committee for the Nobel Prize in Physics?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
4. *Gravitational Wave Detection Facilities.* See the background information on gravity waves discussed in the previous problem. Building facilities to detect these waves will cost more than a quarter of a billion U.S. dollars. Should taxpayer dollars be spent on finding these waves, or is the money better spent elsewhere?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
5. *Falling into a Black Hole.* As shown in this chapter, a person in a spacesuit is tidally distorted when falling feet first into a black hole. The body is stretched lengthwise and compressed widthwise. In reality, the body would be tidally ripped apart as the gravitational pull at the feet is significantly more than at the head. Should the cartoon be redrawn to show the person being tidally ripped apart? Would this image be appropriate for a textbook? If so, should the textbook be labeled with a warning about the nature of the content (like a parental guidance rating system)? What is the minimum age a student should be to learn this information?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- End States of Stars
- Neutron Star
- Schwarzschild Radius

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 14: Neutron Stars and Pulsars

In *Alice in Wonderland*, the Red Queen says, "Why, sometimes I've believed as many as six impossible things before breakfast." Perhaps one of those was neutron stars. They certainly seem impossible, but in fact they are real, and astronomers have found clear evidence to show what they are like. They are remnants left by the supernova deaths of massive stars and have collapsed down to very small size and very high density. An animation in this lab will show you why that collapse leads to very high rates of rotation.

Virtual Astronomy Lab 14, "Neutron Stars and Pulsars," contains an exercise that will show you how a spinning neutron star can appear to be emitting pulses. Astronomers call such objects pulsars, and the timing of the pulses provides important evidence about the pulsars. For one thing, the pulses are very short. Notice in an animation how an object cannot change its brightness abruptly because it takes some time for light to cross its diameter. The short pulses from pulsars were the first clue that pulsars were very small—only an object as small as a neutron star could explain the pulse properties.

As you learn more about pulsars in Lab 14, pay careful attention to the distribution of pulsar periods in the graph. Notice that the fastest pulsars tend to be the oldest and have the weakest magnetic fields. They have been spun up in binary systems, and their weak magnetic fields don't slow them down much. But also notice that the pulsars that emit more than typical amounts of X-rays and gamma rays are young and have powerful magnetic fields. This is the kind of evidence that astronomers use to understand the "back stories" of neutron stars and the life cycles of stars.

Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0, Lab 14, "Neutron Stars and Pulsars."

Virtual Astronomy Lab 15: General Relativity and Black Holes

You can play tennis without thinking about general relativity because a tennis court is small, velocities are low, and the strength of Earth's gravity is not extreme. In that situation, you experience space-time in a very limited way, and Newton's theory of gravity is plenty good enough. But Newton's theory of gravity is not really how nature works. After almost a century of tests, scientists are confident that Einstein's theory of general relativity describes how gravity and space-time really work.

On an Earth-bound tennis court, Einstein's correct predictions about how the tennis ball will behave under the influence of gravity will be so similar to Newton's incorrect predictions that the best measuring equipment available will not be able to tell the difference. And Newton's theory is mathematically much simpler, so you can go with that. But if you could play some cosmic tennis game taking place across cosmological distances, or at velocities near the speed of light, or near very strong gravitational fields such as around a black hole, or all of the above, Einstein's general relativity would tell you the truth about how to play tennis under those conditions and Newton's gravitation equations would not. Einstein's general relativity predicts that the structure of space-time can actually bend, and that produces very strange effects. Some of those effects can actually be detected, and they not only confirm that general relativity is a good description of nature, but they also reveal more about the universe.

Because a strong gravitational field can bend space-time you can detect the effects of gravitational lenses as light follows the bends in space-time near massive objects such as galaxy clusters. Astronomers can measure that apparent displacement of distant objects, and that provides critical information, especially about dark matter, that you will read about in later chapters. The amount of bending of space-time near orbiting neutron stars is constantly fluctuating, and those fluctuations spread outward through space in the form of gravity waves moving at the speed of light. Einstein's predictions of how the orbit of a binary neutron star should evolve as gravity waves carry away the system's energy have been confirmed to impressively high precision. Detecting gravity waves directly is very difficult and has not yet been accomplished, but the astronomers who indirectly confirmed the existence of gravity waves by detecting their effects in a binary pulsar system, Joseph Taylor and Russell Hulse, received the Nobel Prize for their accomplishment.

Virtual Astronomy Lab 15, "General Relativity and Black Holes," leads you through exercises that illustrate the physics of black holes and the highly curved space-time near them. You can arrive at some understanding of concepts such as the *Schwarzschild radius* and the *photon sphere*. Also, you can learn more about how the tides near black holes distort nearby objects and how in-falling matter swirls around a black hole to form a hot accretion disk that can emit X-rays. Astronomers can detect those X-rays and use them to find the properties of black holes. As you work on this lab and look at its animations and exercises, be careful to distinguish among theory, predictions, and observations. You should keep track of what can be observed and what astronomers might learn from those observations. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

17

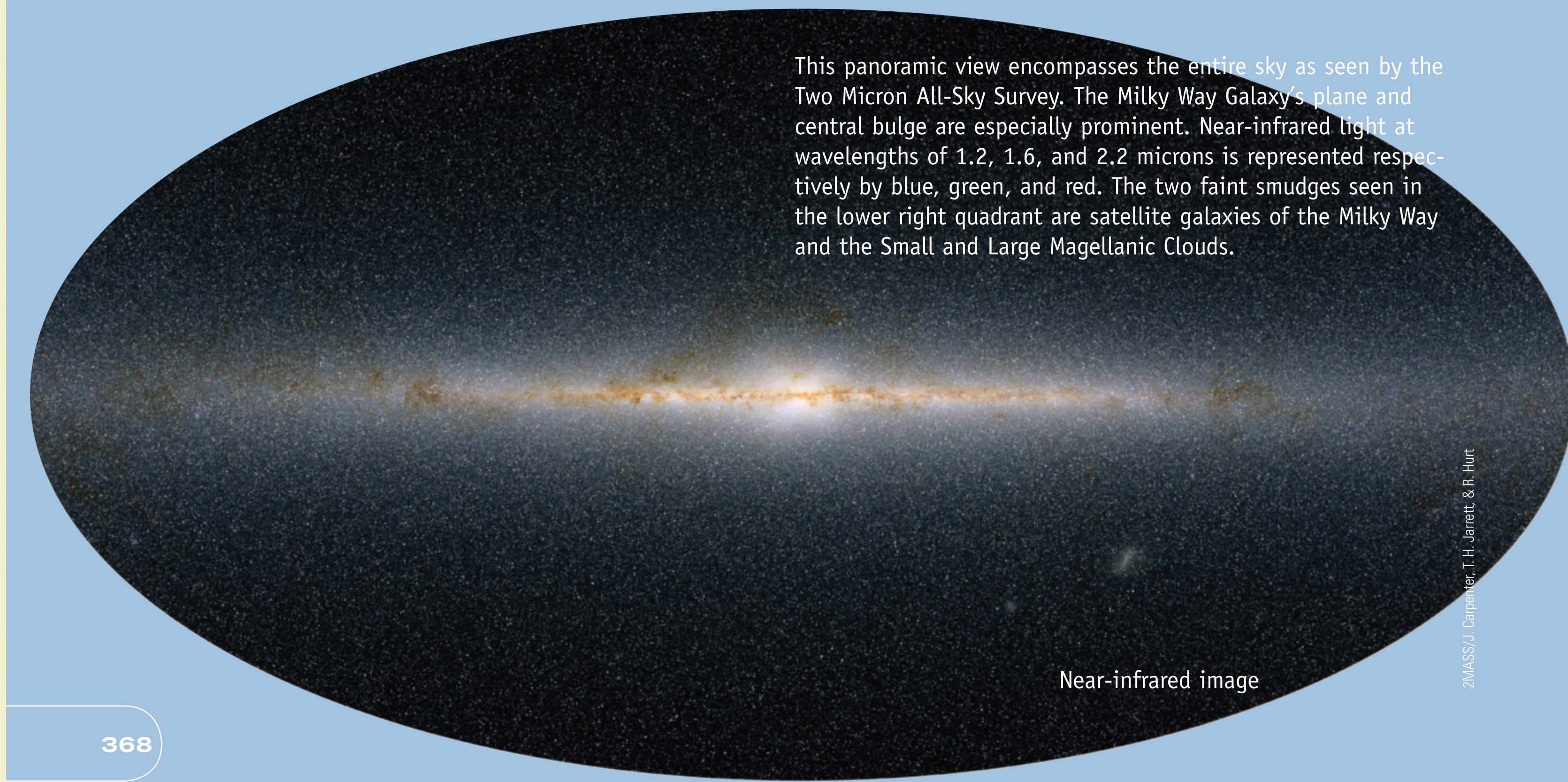
The Milky Way Galaxy

Guidepost

You have traced the life stories of stars from their birth in clouds of gas and dust to their deaths as white dwarfs, neutron stars, or black holes. Now you are ready to step back and view stars in vast communities called galaxies. This chapter focuses on our home galaxy, the Milky Way, and addresses four important questions:

- **What is the evidence that we live in a galaxy?**
- **What is the evidence that our Milky Way is a spiral galaxy, and what are the spiral arms?**
- **What lies at the center of the Milky Way Galaxy?**
- **How did the Milky Way Galaxy form and evolve?**

Discovering the Milky Way Galaxy by answering these questions will be one more step toward understanding the universe as a whole. In the chapters that follow, you will leave your home galaxy and voyage out among the billions of other galaxies that fill the depths of the universe.



This panoramic view encompasses the entire sky as seen by the Two Micron All-Sky Survey. The Milky Way Galaxy's plane and central bulge are especially prominent. Near-infrared light at wavelengths of 1.2, 1.6, and 2.2 microns is represented respectively by blue, green, and red. The two faint smudges seen in the lower right quadrant are satellite galaxies of the Milky Way and the Small and Large Magellanic Clouds.

2MASS/J. Carpenter, T. H. Jarrett, & R. Hurt

The Stars Are Yours.

JAMES PICKERING

THE *STARS ARE YOURS* is the title of a popular astronomy book written by James Pickering in 1948. The point of the title is that the stars belong to everyone equally, and you can enjoy the stars as if you owned them.

Next time you admire the night sky, recall that every star you see is part of the star system in which you live. You will learn in this chapter how evidence reveals that you are inside a great wheel of stars, a galaxy. The Milky Way Galaxy is over 80,000 ly in diameter and contains more than 100 billion stars. It is your galaxy because you live in it, but you are also a product of it, because the stars in the Milky Way Galaxy made many of the atoms in your body. You begin this chapter by pondering the notion that the stars belong to you, but when you reach the end of the chapter, you may realize that you also belong to the stars.

17-1 Discovery of the Galaxy

IT SEEMS ODD to say that astronomers discovered something that is all around us, but it isn't obvious that we live in a galaxy. You might ask, "How do we know what our galaxy is like? Nobody has ever seen it from the outside." Finding the evidence to answer that question was one of the greatest adventures in astronomy.

The Great Star System

Since ancient times, humanity has been aware of a hazy band of light around the sky (■ Figure 17-1). The ancient Greeks named that band *galaxies kuklos,* the "milky circle." The Romans changed the name to *via lactia,* meaning "milky road" or "milky way." It was not until Galileo used his telescope in 1610 that anyone realized the Milky Way is made of a huge number of stars.

Almost every celestial object you can see with your unaided eyes is part of the Milky Way Galaxy. The only exception viewable from Earth's Northern Hemisphere is the Andromeda Galaxy (also known as Messier 31, abbreviated M31), which is barely visible as a faint patch of light in the constellation Andromeda.* You will learn later in this chapter that our galaxy is also a spiral galaxy and, seen from a distance, would look somewhat like the Andromeda Galaxy (Figure 17-1b).

*Consult the star charts at the end of this book to locate the Milky Way and the Andromeda Galaxy, labeled M31 on the chart.

Galileo's telescope revealed that the glowing Milky Way is made up of stars, and later astronomers realized that the sun must be located in a great wheel-shaped cloud of stars, which they called the star system. That star system appears from our location inside it as the band of the Milky Way encircling the sky. In 1750, Thomas Wright, drawing on the technology of the time, referred to the wheel-shaped star system as the grindstone universe, by analogy with the thick disks of stone used in mills. Wright used the term *universe* because, so far as was known at the time, the Milky Way star system was the entire universe.

In the late 18th century astronomers Sir William Herschel and Caroline Herschel (Sir William's sister) set out to map the three-dimensional distribution of stars in the Milky Way. They assumed that they could see to the outer boundaries of the Milky Way in all directions and hypothesized that by counting the number of stars that were visible in different directions they could find the relative distances to the edges of the star distribution. If they saw many stars in one direction, they reasoned that the edge of the Milky Way in that direction is far away. If their telescope revealed fewer stars in another direction, they concluded that edge must be closer. Calling their method "star gauging," they counted stars in 683 directions in the sky and outlined a model of the star system (■ Figure 17-2). Their data indicated that the stars were arranged in a disk shape with the sun near the center. In some directions the Herschels saw very few stars, and these "holes in the sky" produced great irregularities along the edge of their diagram.

The model proposed by the Herschels was widely accepted and studied by other astronomers. The Herschels were not able to measure the size of the star system, but later researchers attempted to do so. By the early 20th century, astronomers had concluded that the star system's disk is about 10 kiloparsecs in diameter and 2 kiloparsecs thick. A **kiloparsec (kpc)** is 1000 pc.

The Herschels assumed that they could see to the edge of the star system, but that was not correct. They were actually seeing only as far into the Milky Way as the interstellar dust permitted. At the time the Herschels did their work, astronomers did not understand that the interstellar medium partially blocks the passage of light (look back to Chapter 14). Because the Herschels counted roughly similar numbers of stars in most directions around the Milky Way, they incorrectly concluded that the sun is near the center of the star system. Furthermore, the "holes in the sky" observed by the Herschels are not empty but are now known to be especially dense interstellar clouds completely blocking the view of stars beyond them.

Modern astronomers know that the Milky Way Galaxy is much larger than was first thought, and that the sun is not at its center. How the human race realized the truth about our location in the universe is a story that begins with a woman

Figure 17-1

(a) Nearby stars look bright, and people around the world group them into different constellations. Nevertheless, the vast majority of the stars in our galaxy that are visible from Earth merge into a faintly luminous path that circles the sky, the Milky Way. This artwork shows the location of a portion of the Milky Way near a few bright winter constellations. (See Figure 17-5 and the star charts at the end of this book to locate other sections of the Milky Way in your sky.) (b) The Andromeda Galaxy, a spiral galaxy about 2.5 million ly from Earth, looks somewhat as our galaxy would appear if you could view it from the same distance.

studying stars that pulsate and leads to a man observing distant star clusters.

Variable Stars and Distance Estimates

It is a **Common Misconception** that the stars are eternal and unchanging, but astronomers have known for centuries that some stars change in brightness. Of course, novae and supernovae appear, grow brighter, and then fade, but many stars change periodically, growing brighter, then fainter, then brighter, and so on. Some of these variable stars are eclipsing binaries, but some are stars that pulsate like beating hearts. The period of pulsation is the time it takes a star to complete a cycle from bright to faint to bright again.

In 1912, Henrietta Leavitt was studying a star cloud in the southern sky known as the Small Magellanic Cloud. On her photographic plates, she found many variable stars, and she noticed that the brightest had the longest periods. She didn't know the distance to the cloud, so she couldn't calculate absolute magnitudes (look back to Chapter 13), but because all of the variables she was observing are in the same star cloud and can be assumed to be at about the same distance, she concluded that there is a relationship between period and luminosity.

The variable stars Leavitt saw, **Cepheids,** are named after the first such star discovered, δ Cephei. Many Cepheid variables are known today; they have pulsation periods from 1 to 60 days and lie in a region of the H–R diagram known as the **instability strip**

(■ Figure 17-3). Cepheids are clearly giant and supergiant stars. A related kind of variable, **RR Lyrae stars,** are named after the variable star RR in the constellation Lyra. They have a pulsation period of about half a day and lie at the low-luminosity end of the instability strip.

Stars in the instability strip are unstable and pulsate because a layer of partially ionized helium in the star's atmosphere can absorb and release energy like a spring. In very hot stars, the layer is too high to make the star unstable, and in very cool stars, the layer is too low. Stars in the instability strip happen to have this layer in just the right place to make them pulsate as energy flows out of their interiors. As stars evolve across the H–R diagram into the instability strip, they become unstable; they stop pulsating when they evolve out of the strip.

Massive stars are very luminous, and they cross the instability strip higher in the H–R diagram. Because these massive stars are larger, they pulsate slower, just as large bells vibrate slower and have deeper tones. Lower-mass stars are less luminous, cross the instability strip lower in the H–R diagram, and, being smaller, pulsate faster. This explains why the long-period Cepheids are more luminous than the short-period Cepheids. Leavitt first noticed this in 1912, and it is now known as the **period-luminosity relation** (■ Figure 17-4). You will remember from Chapter 13 that Favorite Star Polaris, the North Star, is a supergiant; it lies high in the instability strip and pulsates as a Cepheid variable.

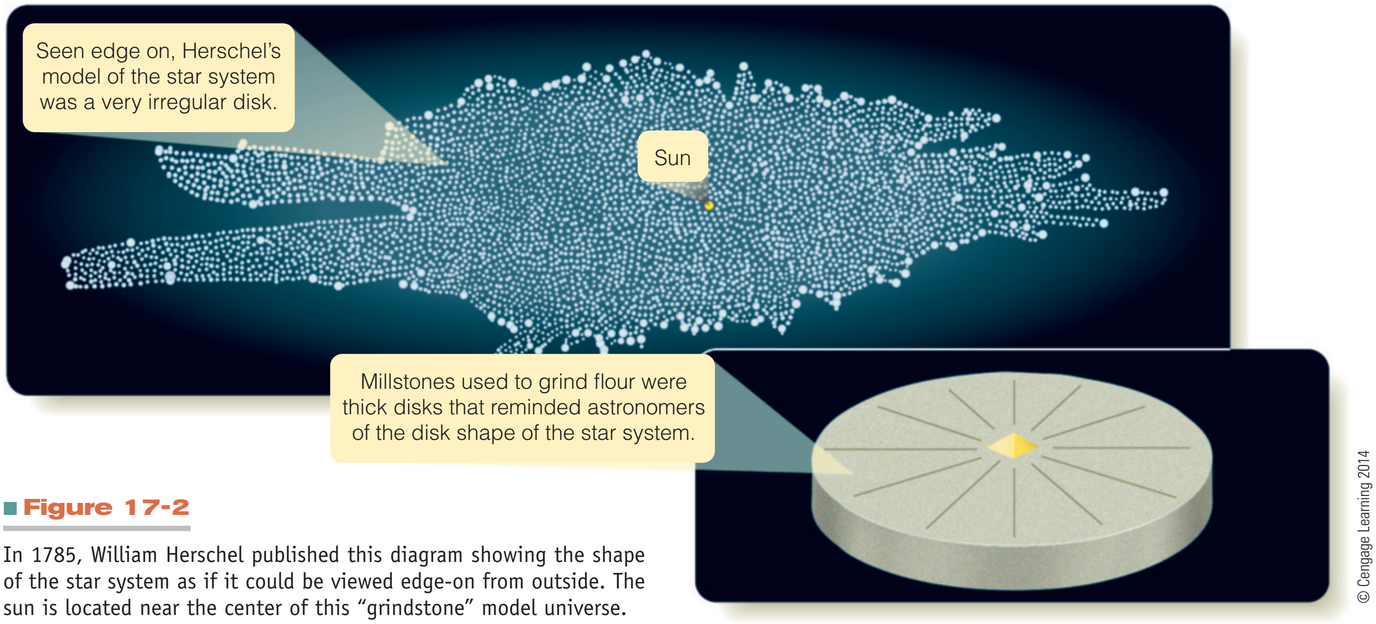
The physics of Cepheid pulsating stars was unknown in 1912. Although Leavitt didn't know the distance to the Small Magellanic Cloud, she knew their periods of pulsations are related to their luminosities. Finding those luminosities was the key that unlocked the secret of the Milky Way's size.

Size of the Galaxy

Until early in the 20th century no one knew that humanity lives inside a galaxy. That began to change around 1920 when an astronomer named Harlow Shapley discovered how big our star system really is. Besides being one of the turning points of modern astronomy, Shapley's study illustrates one of the most common techniques in astronomy. If you want to understand some of the basic ways astronomers find out about the universe, then Shapley's part of the story is worth following in detail.

You learned in Chapter 15 (pages 323–324) that star clusters come in two quite different types, open clusters and globular clusters. Shapley began his research on the star system by noticing that, although open star clusters are scattered all along the Milky Way, more than half of all globular clusters lie in or near the constellation Sagittarius. The globular clusters seem to be distributed in a great cloud with a center located in the direction of Sagittarius (■ Figure 17-5). Shapley assumed that the orbital motion of these clusters is controlled by the gravitation of the entire star system, and for that reason he concluded that the center of the star system could not be near the sun but must lie somewhere toward Sagittarius.

To find the distance to the center of the star system and thus the size of the star system as a whole, Shapley needed to find the distance to the star clusters, but that was difficult to do. The clusters are much too far away to have measurable parallaxes.



■ **Figure 17-2**

In 1785, William Herschel published this diagram showing the shape of the star system as if it could be viewed edge-on from outside. The sun is located near the center of this "grindstone" model universe.

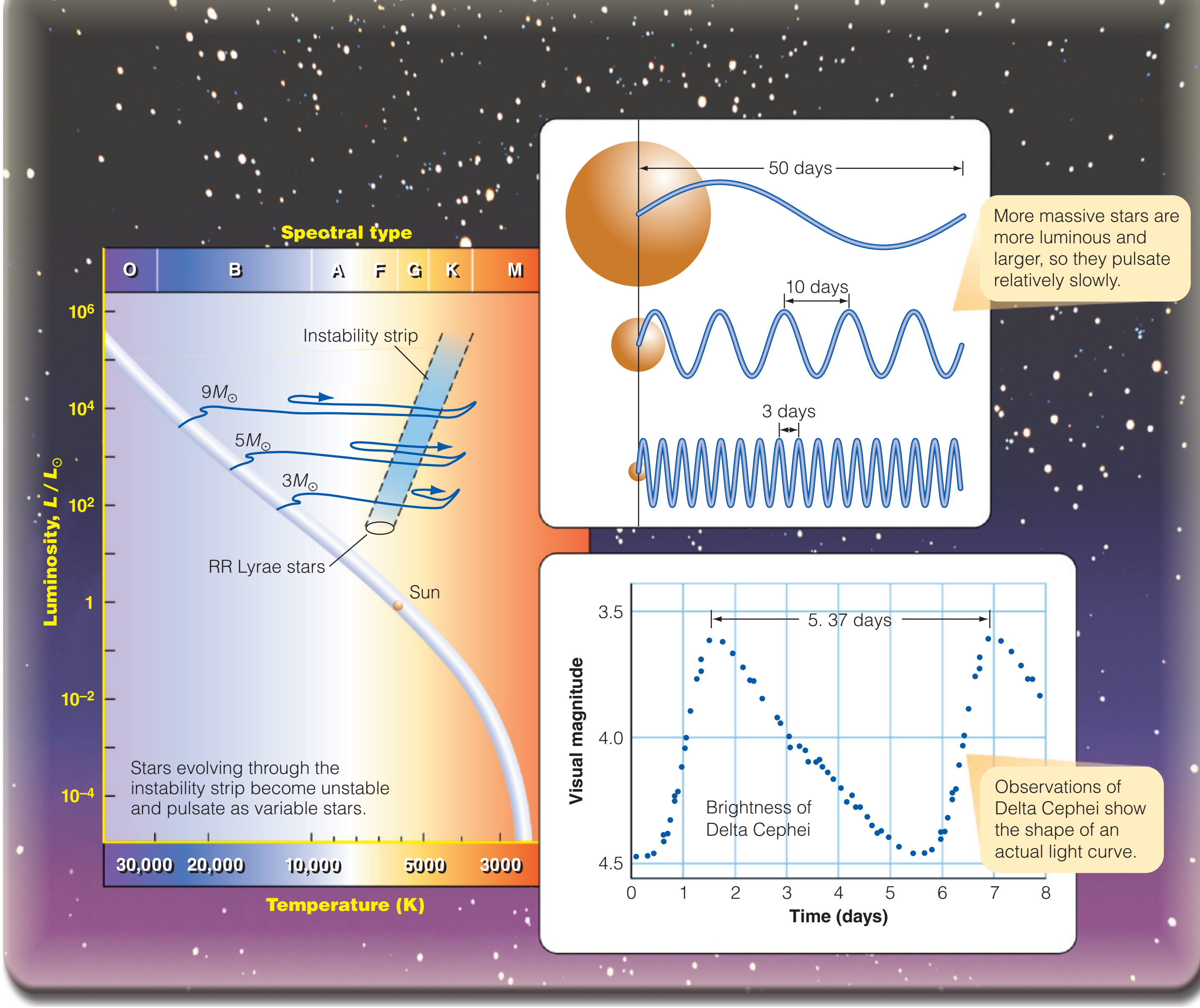


Figure 17-3

The instability strip on the H–R diagram contains combinations of stellar temperatures and luminosities corresponding with unstable, pulsating internal structures. The more massive a star is, the more luminous it is, and the larger in diameter it becomes when it leaves the main sequence. Those larger stars pulsate with longer periods when they pass through the instability strip. Therefore, because both luminosity and period of pulsation depend on mass, there is a relationship between period and luminosity.

They do, however, contain variable stars, and those were the beacons Shapley needed to find the distances to the clusters.

Shapley knew about Leavitt's work on Cepheid variable stars. Shapely and other astronomers realized that the Cepheids could be used to determine distances if their true luminosities could be discovered, but because Cepheids are giant and supergiant stars, they are relatively rare. None lies close enough to Earth to have a measurable parallax, but their **proper motions**—apparent slow movements across the sky due to their motions through space—can be measured. The more distant a star is, the smaller its proper motion tends to be, so proper motions contain clues to distance. Shapley found 11 Cepheids with measured proper motions and used a statistical process to find their average distance and thus their average absolute magnitude. Then, he could replace Leavitt's *apparent* magnitudes on the vertical axis of the period-luminosity diagram (Figure 17-4) with *absolute* magnitudes, representing true luminosities. All of the stars in the diagram were thus **calibrated** as standard light sources that astronomers could use to find distances.

Figure 17-4

The period–luminosity diagram is a graph of the brightness of variable stars versus their periods of pulsation. You could plot brightness as luminosity, but because the diagram is used in distance calculations, it is more convenient with absolute magnitude on the vertical axis. Modern astronomers know that there are two types of Cepheids, something that astronomers in the early 20th century could not recognize in their limited data.

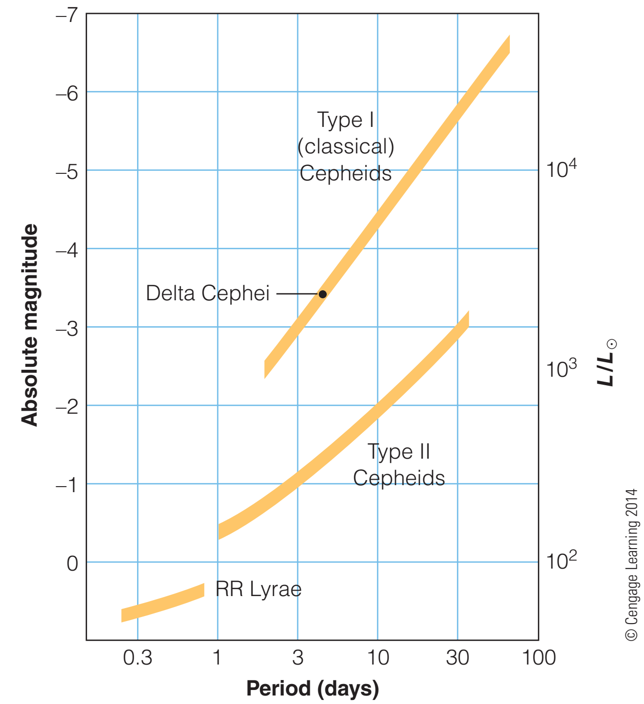
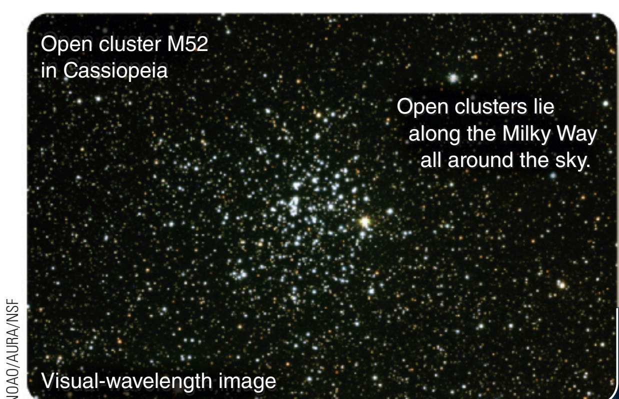
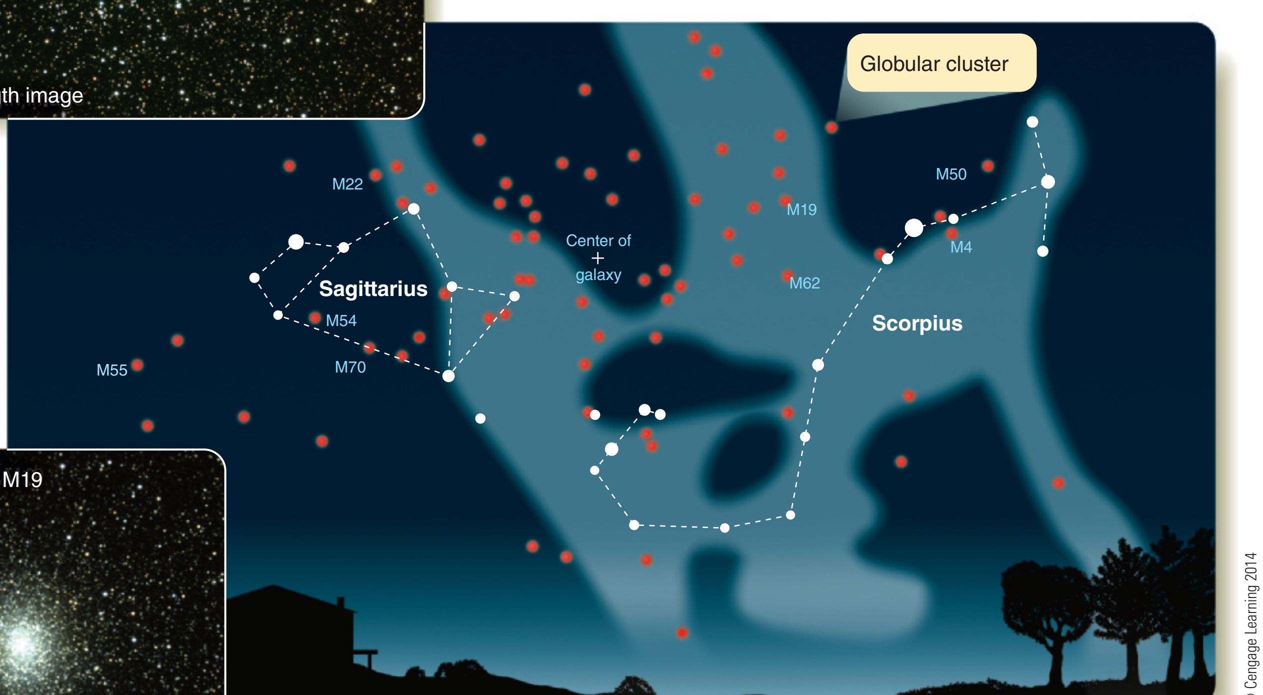
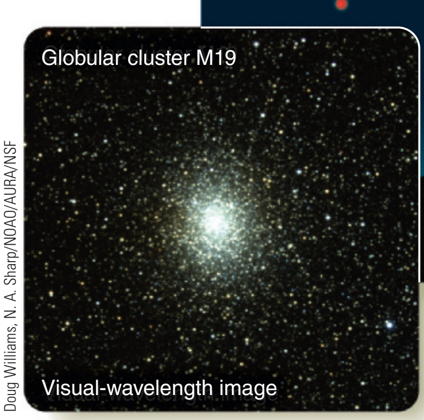


Figure 17-5

Nearly half of cataloged globular clusters (red dots) are located in or near Sagittarius and Scorpius. A few of the brighter globular clusters, labeled with their catalog designations, are visible in binoculars or small telescopes. Constellations are shown as they appear above the southern horizon on a summer night as seen from latitude 40° N, typical for most of the United States.







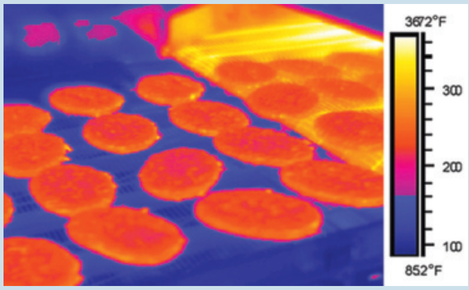
Globular clusters are scattered over the entire sky but are strongly concentrated toward Sagittarius.

Calibration

How do you take the temperature of a vat of molten steel? Astronomers often say that Shapley "calibrated" the Cepheids for the determination of distance, meaning that he did all the detailed background work to determine the luminosities of Cepheids. After that, he and other astronomers could use his calibrated period-luminosity diagram to find the distance to other Cepheids without repeating the detailed calibration.

Calibration is actually very common in science because it saves a lot of time and effort. For example, engineers in steel mills must monitor the temperature of molten steel, but they can't dip in a thermometer. Instead, they can use handheld devices that measure the color of molten steel. You recall from Chapter 6 that the color of blackbody radiation is determined by the temperature of the body emitting it. Molten steel emits visible and infrared radiation that is nearly perfect blackbody radiation, so the manufacturer can calibrate the engineer's devices to convert the measured color to a temperature displayed on digital readouts. The engineers don't have to repeat the calibration every time; they just point their instrument at the molten steel and read off the temperature. Astronomers have made the same kind of color-temperature calibration for stars.

As you read about any science, notice how calibrations are used to simplify common measurements. But notice, too, how important it is to get the calibration right. An error in calibration can throw off every measurement made with that calibration.



Courtesy of FLIR Systems, Inc.

An infrared video camera calibrated to measure temperature allows bakers to monitor the operation of their ovens.

Calibrations like this one are important tools in science (**How Do We Know? 17-1**).

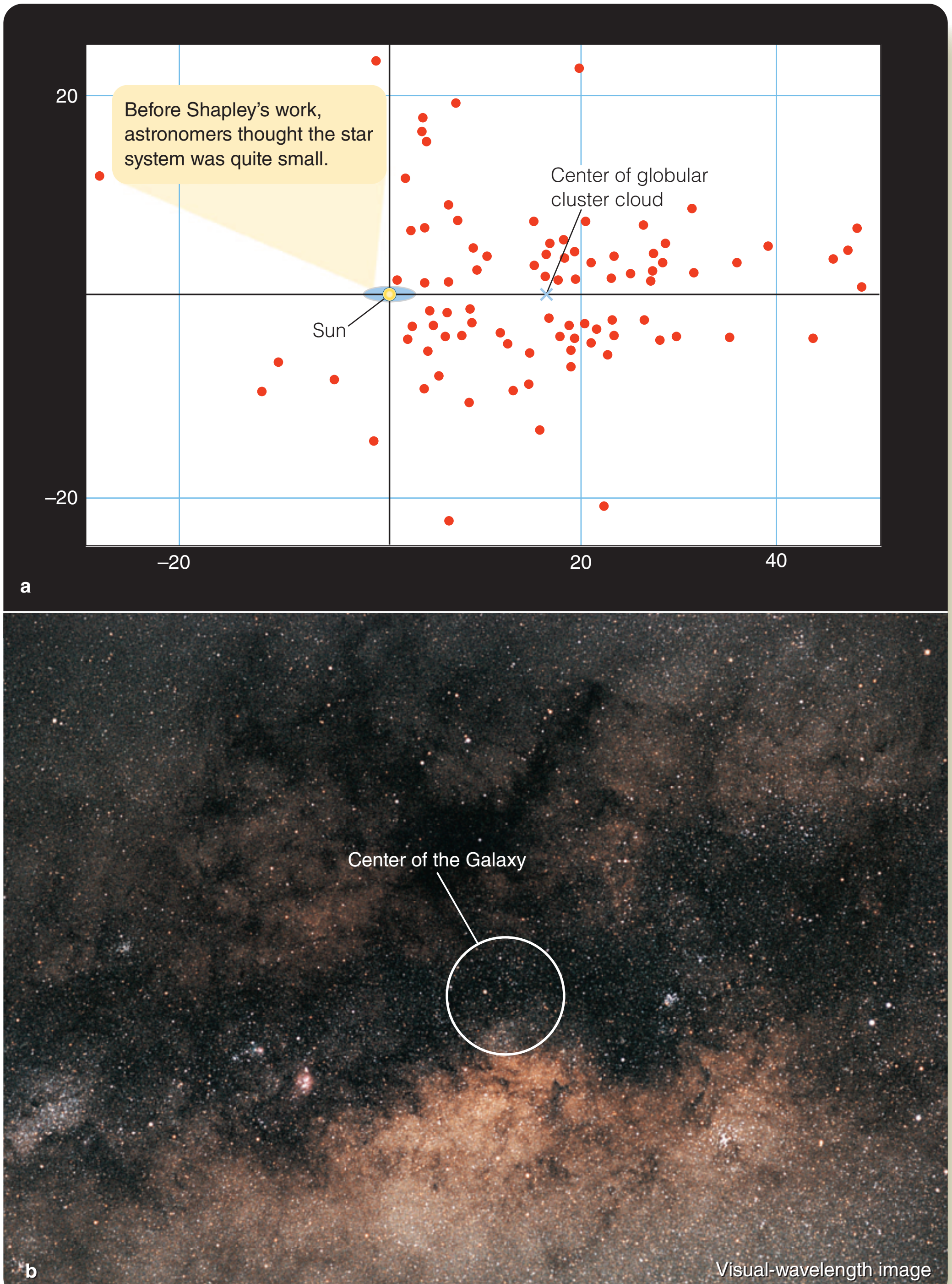
Once Shapley had calibrated the period-luminosity relation, he could use it to find the distance to any cluster in which he could identify those types of variable stars. By taking a series of photographic plates over a number of nights, Shapley was able to pick out the variable stars in a cluster, measure their average apparent magnitude, and find their periods of pulsation. Knowing that, he could read their absolute magnitude off the period-luminosity diagram. With the apparent magnitude and the absolute magnitude, he could calculate the distance to the star cluster (look back to Reasoning with Numbers 13-2, page 263).

Shapley later wrote that it was late at night when he finally plotted the directions and distances to the globular clusters on graph paper and found that, just as he had supposed, they formed a great swarm whose center lay many thousands of light-years in the direction of Sagittarius, confirming his suspicion that the center of the star system was not near the sun but was far away in Sagittarius (■ Figure 17-6). He found the only other person in the building, a cleaning lady, and the two stood looking at his graph as he explained that they were the only two people on Earth who understood that humanity lives, not near the center of a small star system, but in the suburbs of a vast wheel of stars.

It is important to note that there were problems with Shapley's calibrations and distance determinations. Shapley was able to observe the globular clusters at great distances because they lie outside the plane of the galaxy (Figure 17-6) and are not dimmed much by the interstellar medium, but the Cepheid stars he used for calibration mostly lie in the plane of the galaxy and were strongly affected by extinction. Also, Shapley did not know there are several types of variable stars, but the ones he used for the distance estimates were different from the ones he used for the calibration. Consequently his estimate for the size of the galaxy shown in Figure 17-6 was bigger than the modern value. Nevertheless, he got the main point right, that we are thousands of light-years from the center of the galaxy.

Building on Shapley's work, other astronomers began to suspect that some of the faint patches of light visible through telescopes were other star systems. Within a few years they found evidence that the faint patches of light were indeed other galaxies much like our own Milky Way Galaxy. Today the largest telescopes can detect an estimated 100 billion galaxies similar to our own. Our home galaxy is special only in that it is our home.

Shapley's study of star clusters led to the discovery that we live in a galaxy and that the universe is filled with similar galaxies. Like Copernicus, Shapley moved us from the center to the outskirts. A careful analysis of the structure of the galaxy will tell you the details about our real location.



© Cengage Learning 2014

Daniel Good

Figure 17-6

(a) Shapley's study of globular clusters showed that they were not centered on the sun, at the origin of this graph, but rather formed a great cloud centered far away in the direction of Sagittarius. Distances on this graph are given in thousands of parsecs, corresponding to Shapley's original calibration that produced values more than 2 times larger than the modern calibration. (b) Looking toward Sagittarius at visual wavelengths, you see nothing to suggest that this is the center of the galaxy. Interstellar dust and gas block your view. Only the distribution of globular clusters told Shapely that the center lies in this direction.

17-2 Structure of the Galaxy

Astronomers commonly give the diameter of the Milky Way Galaxy as approximately 25 kpc, or 80,000 ly, and place the sun about two thirds of the way from the center to the edge (Figure 17-7). Note that this is only the diameter of the easily detected part of our galaxy. Later in this chapter you will learn that the entire galaxy, some parts of which are not easy to observe, is much larger than 25 kpc in diameter.

Components of the Galaxy

The disk of the galaxy is often referred to as the **disk component.** It contains most of the galaxy's stars and nearly all of its gas and dust. Beca use the disk is home to the giant molecular clouds within which stars form (look back again to Chapter 14), nearly all star formation in our galaxy takes place in the disk. Most of the stars in the disk are middle- to lower-main-sequence stars like the sun, a few are red giants and white dwarfs, and fewer still are brilliant blue O and B stars. Those hot, massive stars are rare, but they are so luminous that they provide much of the light from the disk.

Although astronomers can survey the locations and distances of the stars, they cannot cite a single number for the thickness of the disk because it lacks sharp boundaries. Also, the thickness of the disk depends on the kind of object studied. Stars like the sun, with ages of a few billion years, lie within about 500 pc above

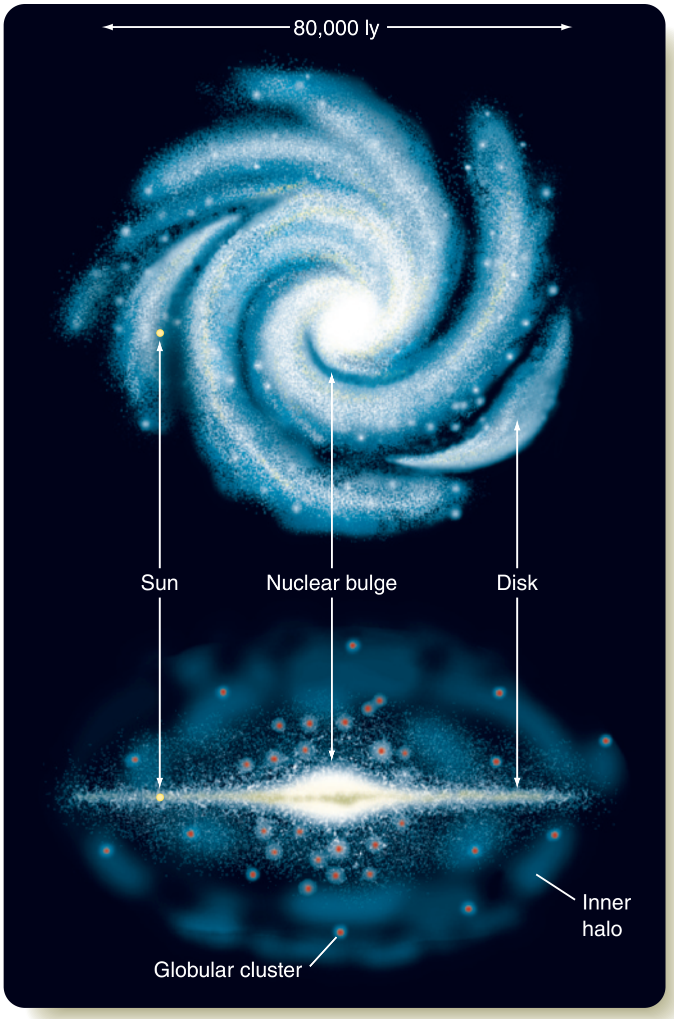


Figure 17-7

Sketches of our Milky Way Galaxy viewed face-on and edge-on. Note the position of the sun and the distribution of globular clusters in the halo. Hot blue stars light up the spiral arms. Only the inner halo is shown here. At this scale, the entire halo would be larger than a dinner plate.

and below the central plane. In comparison, the youngest stars, including the O and B stars and the gas and dust from which stars form, are confined to a thin disk extending only about 50 pc above and below the plane (■ Figure 17-8). In proportion to its diameter of about 25 kpc, the disk of the galaxy is thinner than a thin pizza crust.

If you look "up" or "down," out of the galaxy's disk, you are looking away from the dust and gas, so you can see into the galaxy's **halo,** a spherical cloud of stars and star clusters that contains almost no gas and dust. Because the halo contains no dense gas clouds, it cannot make new stars. Halo stars are old, cool, lower-main-sequence stars, red giants, and white dwarfs. It is difficult to judge the extent of the halo, but it could have as much as 10 times the diameter of the visible disk.

Around the center of our galaxy lies the **central bulge,** a flattened cloud of billions of stars about 6 kpc in diameter (Figure 17-7). Like the halo, it contains little gas and dust. Astronomers often refer to the halo and the central bulge together as the **spherical component** of the galaxy. The central bulge is the most crowded part of the spherical component. The central bulge, visible above and below the plane of the galaxy and through gaps in the obscuring dust, consists mostly of stars that are old and cool like the stars in the halo.

The halo includes approximately 200 globular clusters, each of which contains 50,000 to a million stars in a sphere a few tens of parsecs in diameter (Figure 17-5). You can tell their average age is about 11 billion years from the positions of the turnoff points in their H–R diagrams.

Orbital motions are quite different in the two components of the galaxy. Disk stars follow nearly circular orbits that lie in the plane of the galaxy (■ Figure 17-9a). Halo stars and globular clusters, however, follow highly elongated orbits tipped steeply to the plane of the disk (Figure 17-9b). (Although the diagram shows these orbits as elliptical, that is not quite true. The gravitational influence of the thick bulge forces them into curves that do not quite return to the same starting point.) The dramatic difference between the motions of halo stars and disk stars will be important evidence when you consider the formation of the galaxy later in this chapter.

This analysis of the components of the Milky Way Galaxy may leave you wondering about one important question: How much matter does our galaxy contain? To answer that question, you can measure the galaxy's rotation.

Mass of the Galaxy

To find the mass of an object, astronomers must observe the motion of another object orbiting it, just as in a binary star system. Humans don't live long enough to see stars move significantly along their orbits around the galaxy, but astronomers can observe the radial velocities, proper motions, and distances of stars and then calculate the sizes and periods of their orbits. The results can reveal the mass of the galaxy.

Stars in the disk of the galaxy follow nearly circular orbits that lie in the plane of the disk (look again at Figure 17-9a). Current observations indicate that the sun orbits the center of our galaxy at about 240 km/s, moving in the direction of Cygnus. The evidence suggests the sun's orbit is nearly circular, so given the current best estimate for the distance to the center of our galaxy, 8.2 kpc, you can find the circumference of the sun's orbit by multiplying by 2π. If you divide the circumference of its orbit by its orbital velocity, you will discover that the sun has an orbital period of about 210 million years.

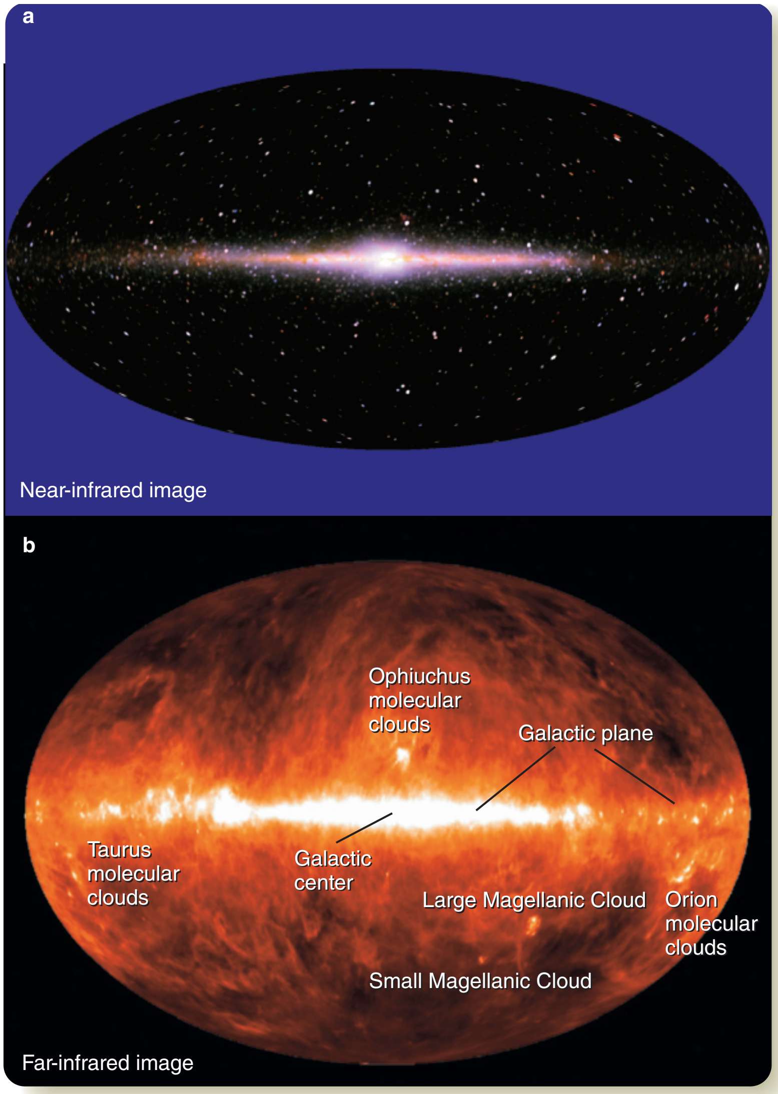
When you studied binary stars in Chapter 13, you learned that the total mass (in solar masses) of a binary star system equals the cube of the separation of the stars a (in AU) divided by the

(110 billion) $M_{\odot}$. This is only a rough estimate because it does not include any mass lying outside the orbit of the sun. Allowing for that overlooked mass yields a lower limit to the total mass for our galaxy of at least 4×10^{11} (400 billion) solar masses

The rotation of our galaxy is actually the orbital motion of each of its stars around the center of mass. Stars at different distances from the center revolve around the center of the galaxy with different periods, so stars starting near each other will draw apart as time passes (■ Figure 17-10), a situation called differential rotation. (Recall that differential rotation was first defined regarding the sun in Chapter 7.)

To fully describe the rotation of our galaxy, astronomers graph orbital velocity versus radius to produce a **rotation curve** (■ Figure 17-11). If all of the mass of the galaxy were concentrated near its center, you would expect to see orbital velocities decrease as you move away from the center. That is what you see in our solar system, where nearly all of the mass is concentrated in the sun. For that reason this is called **Keplerian motion,** referring to Kepler's laws. In contrast, the best observations of the rotation curve of the Milky Way show that orbital velocities in the outer disk are constant or even increasing with distance from the center of the galaxy. Those high orbital velocities indicate that the larger orbits enclose more mass and imply that our galaxy has much more mass than is contained within the radius of the sun's orbit.

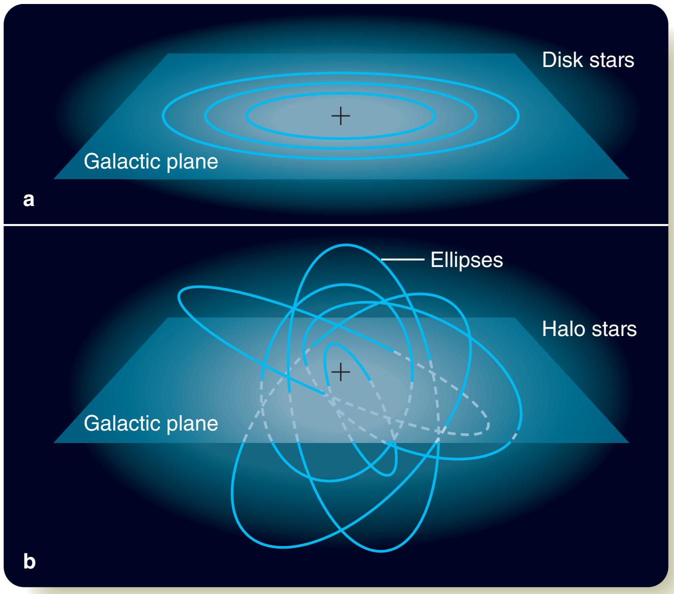
Both observational evidence and mathematical models show that the extra mass lies in an extended halo sometimes called a **galactic corona** that may reach out



■ **Figure 17-8**

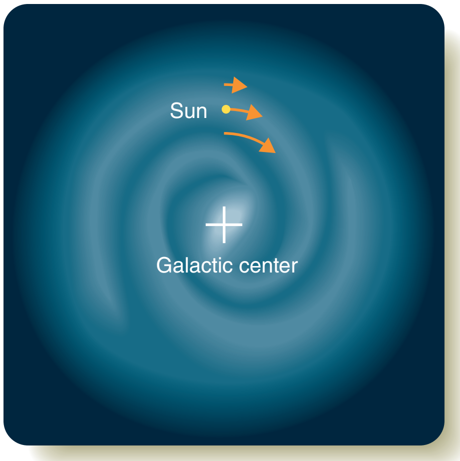
In these infrared images, the entire sky has been projected onto ovals with the center of the galaxy at their centers. The central plane of the Milky Way extends from left to right. (a) At near-infrared wavelengths, the central bulge is prominent, and dust clouds partially block your view. (Compare with the image on page 368 made at approximately the same wavelengths with a larger telescope's higher spatial resolution.) (b) At far-infrared wavelengths, the dust emits blackbody radiation and glows brightly.

square of the period P (in years) (Reasoning with Numbers 13-4). You can use the same procedure here to find the mass of the galaxy. The radius of the sun's orbit around the galaxy is about 8200 pc, and each parsec contains 2.1×10^5 AU. Multiplying, you find that the radius of the sun's orbit is 1.7×10^9 AU. The orbital period is 210 million years, so the mass is 1.1×10^{11}



■ **Figure 17-9**

(a) Stars in the galactic disk have nearly circular orbits that lie in the plane of the galaxy. (b) Stars in the halo have randomly oriented, highly eccentric, elongated orbits.



■ **Figure 17-10**

The differential rotation of the galaxy means that stars at different distances from the center have different orbital periods. In this example, the star just inside the sun's orbit has a shorter period and pulls ahead of the sun, while the star outside falls behind.

■ **Figure 17-11**

The rotation curve of our galaxy is plotted here as orbital velocity versus radius. Data points show measurements made by radio telescopes. Observations outside the orbit of the sun are much more uncertain, and the data points scatter widely. Orbital velocities do not decline outside the orbit of the sun, as you would expect if most of the mass of the galaxy were concentrated toward the center (Keplerian motion). Rather, the curve is approximately flat at great distances, suggesting that the galaxy contains significant mass outside the orbit of the sun.

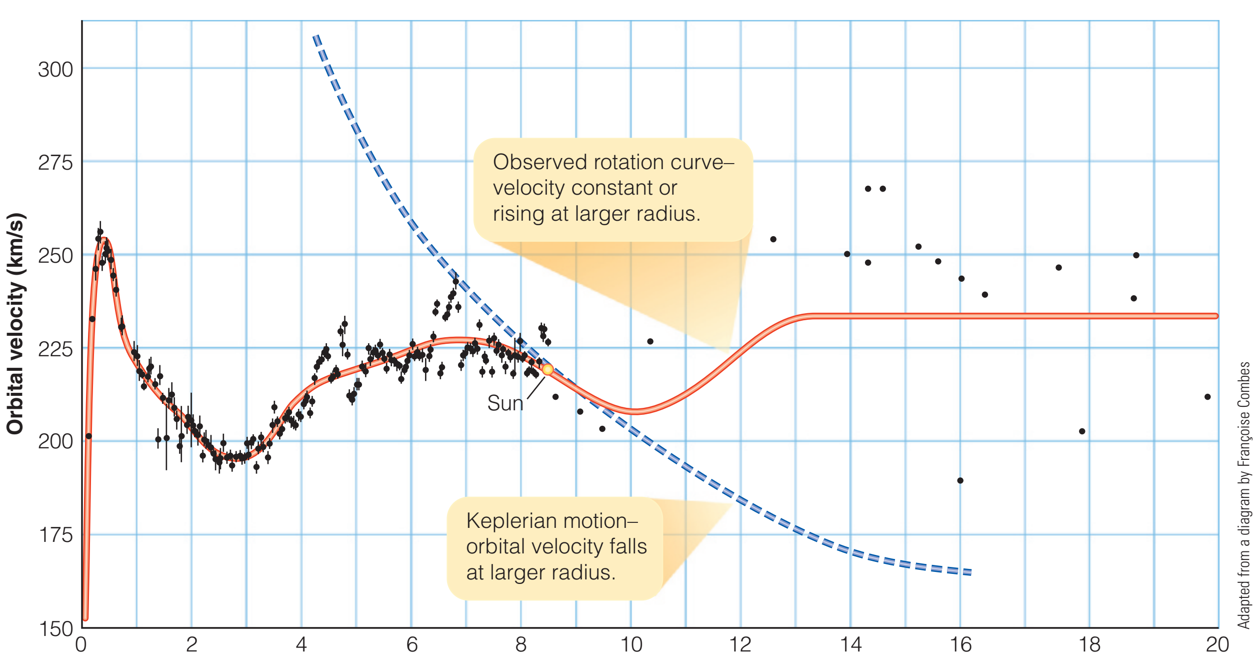
to 10 times farther than the edge of the visible disk and could contain more than a trillion (10^{12}) solar masses. Most of that mass is invisible, neither emitting nor absorbing light, so astronomers refer to it as **dark matter.** Some of the mass in the galactic corona is made up of low-luminosity stars and white dwarfs, but much of the mass must be some other form of matter. You will learn more about the problem of dark matter in the following two chapters. The nature of dark matter is one of the most important unknowns in modern astronomy.

SCIENTIFIC ARGUMENT

If dust and gas block the view, how do astronomers know how big our galaxy is?

Because scientific arguments depend ultimately on evidence, they must explain how scientists know what they know. Interstellar dust in our galaxy blocks the view only in the plane of the galaxy. When you use a telescope to look away from the plane of the galaxy, you are looking out of the obscuring interstellar medium and can see to great distances. The globular clusters are scattered through the halo, with a strong concentration in the direction of Sagittarius. When astronomers look above or below the plane of the galaxy, they can see those clusters, and careful observations reveal variable stars in the clusters. They can then find the distances to those clusters by using modern calibrations of the Cepheid and RR Lyrae variable stars' period-luminosity diagrams. Astronomers infer that the distance to the center of the distribution of globular clusters is the distance to the center of the galaxy because they assume that the distribution of clusters is controlled by the gravitation of the galaxy as a whole.

In any logical argument, it is important to ask "How do we know?" and that is especially true in science. Create a new argument: **What measurements and assumptions reveal the total mass of our galaxy?**



17-3 Spiral Arms and Star Formation

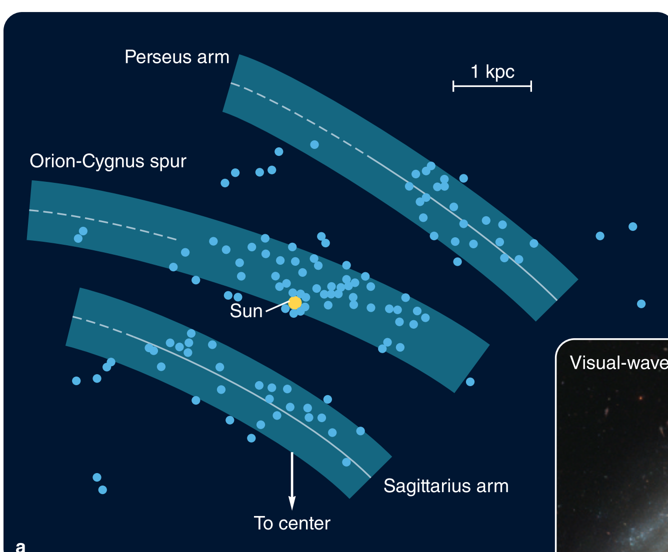
The most striking feature of galaxies like the Milky Way is their patterns of spiral arms that wind outward through the disk. These arms contain swarms of hot, blue stars, clouds of dust and gas, and young star clusters. The young objects suggest that the spiral arms involve star formation, but as you try to understand the spiral arms, you need to consider two problems. First, how can anyone be sure our galaxy has spiral arms if interstellar dust obstructs our view? Second, why doesn't the differential rotation of the galaxy destroy the arms? The solution to both problems involves star formation.

Tracing the Spiral Arms

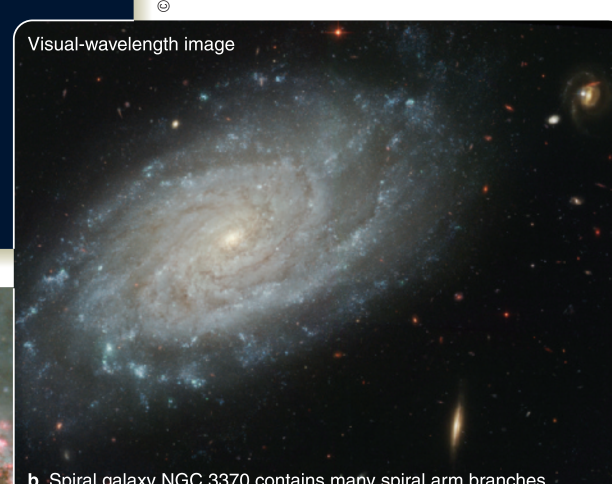
As you have already learned, O and B stars are often found in associations and are very luminous. Thus, they are easy to detect across great distances. Unfortunately, at those great distances their parallaxes are difficult to observe. The available data indicate that OB associations near the sun are not located randomly but lie along several spiral arm segments (■ Figure 17-12a). Those spiral segments have been named for the prominent constellations through which they pass. Astronomers are now using infrared and radio telescopes to penetrate the interstellar dust, locate more distant OB associations, and trace the spiral arms even farther.

Objects used to map spiral arms are called **spiral tracers.** Aside from OB associations, spiral tracers include young open star clusters, clouds of hydrogen ionized by hot stars (emission nebulae), and certain higher-mass variable stars. Notice that all spiral tracers are young objects, formed recently, astronomically speaking. O stars, for example, live for only a few million years. Their typical orbital velocity is about 200 km/s, so they cannot move more than about 500 pc in their lifetimes. This is less than the width of a spiral arm. Because they don't live long enough to move away from the spiral arms, they must have formed there.

Studies of other galaxies show that their spiral arms are also marked by hot blue stars and other spiral tracers (Figures 17-12b

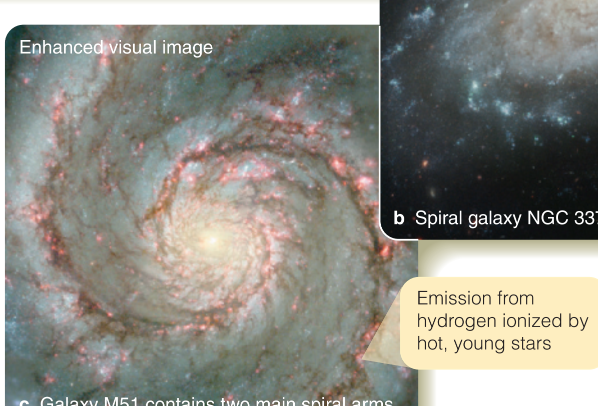


© Cengage Learning 2014



b Spiral galaxy NGC 3370 contains many spiral arm branches.

A. Riess [STScI]



c Galaxy M51 contains two main spiral arms.

NASA and The Hubble Heritage Team [STScI/AURA]

■ **Figure 17-12**

(a) Nearby young O and B stars in our galaxy fall along bands that appear to be segments of spiral arms. (b) Many of the galaxies in the sky are disk shaped like our galaxy, and most of those galaxies have spiral arms marked especially by hot luminous stars and emission nebulae that must be very young. This might cause you to suspect that spiral arms are related to star formation. (c) Images of other galaxies show spiral arms that are especially prominent and well-defined.

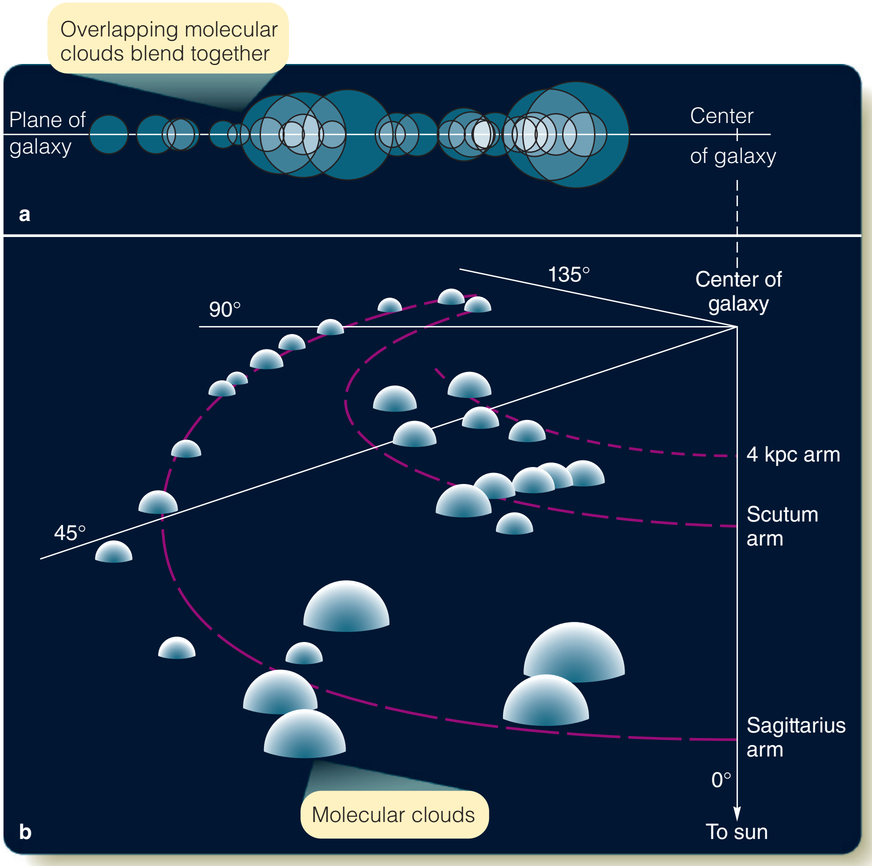


Figure 17-13

(a) At the wavelength emitted by the CO molecule, astronomers find many molecular clouds along the Milky Way, but the clouds overlap in a confusing jumble. By using a model of the rotation of the galaxy to interpret the radial velocities of the clouds, radio astronomers can estimate the distances to each cloud and use them to map spiral arms. (b) This map represents the view from a point 2 kpc directly above the sun. The molecular clouds, shown here as hemispheres extending above the plane of our galaxy, are located along spiral arms. Angle labels show longitudes in a coordinate system with origin at the galactic center.

and 17-12c). The youth of spiral tracers is an important clue about spiral arms. Obviously spiral arms are associated with star formation. Before you can follow this clue, you need to extend your map of spiral arms to show the entire galaxy.

Radio Maps of Spiral Arms

Radio astronomers use the strong spectral line emission from carbon monoxide (CO) to map the location of giant molecular clouds in the plane of the galaxy. (The dust that blocks our view at visual wavelengths is transparent to radio waves because radio wavelengths are much larger than the diameter of the dust particles.) Recall from Chapter 14 that giant molecular clouds are sites of active star formation. If you point a radio telescope at a section of the Milky Way, you will receive a combination of signals from gas clouds in the direction you are looking that lie at various distances across the galaxy. Signals from those different clouds can be separated by measuring the Doppler shifts of their spectral lines. Astronomers then use a simple model of orbital velocities at different distances from the center of the galaxy in order to untangle the observations and locate the individual clouds. Maps constructed from such observations reveal that the giant molecular clouds, like O and B stars, are located along segments of spiral arms (■ Figure 17-13).

As you learned in Chapter 14, observing at a wavelength of 21-cm allows detection of a spectral line of atomic hydrogen. That type of gas is generally warmer than the material in molecular clouds. Just as for the molecular clouds, analysis of the atomic cloud radial velocity data requires that astronomers start with estimates of the orbital velocities at different distances from the center of the galaxy. The atomic hydrogen map (■ Figure 17-14) doesn't trace as clear a spiral pattern as the molecular gas, in part because the warmer atomic gas has more random motions due to its temperature as well as turbulence within the

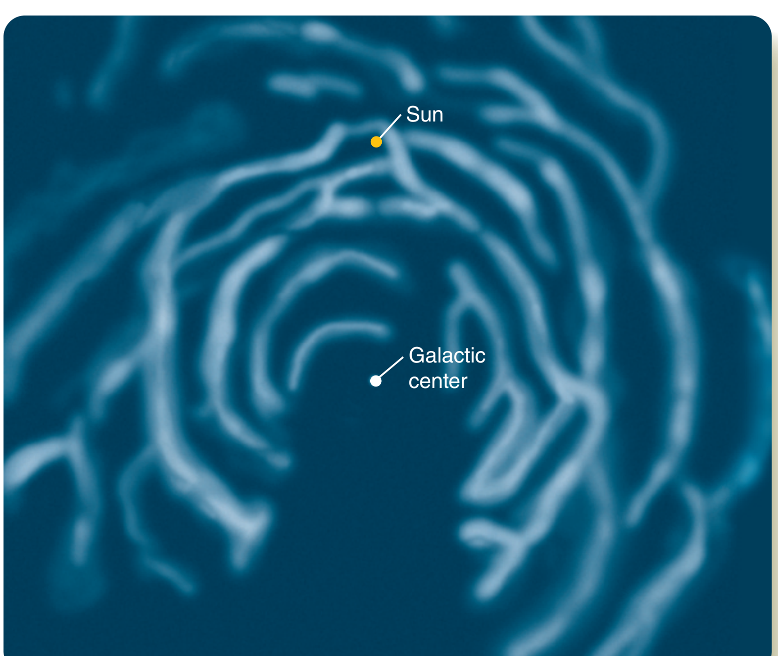


Figure 17-14

This 21-cm radio map of our galaxy confirms that concentrations of neutral atomic hydrogen gas lie approximately in a spiral, but the pattern is complex and suggests branches and spurs.

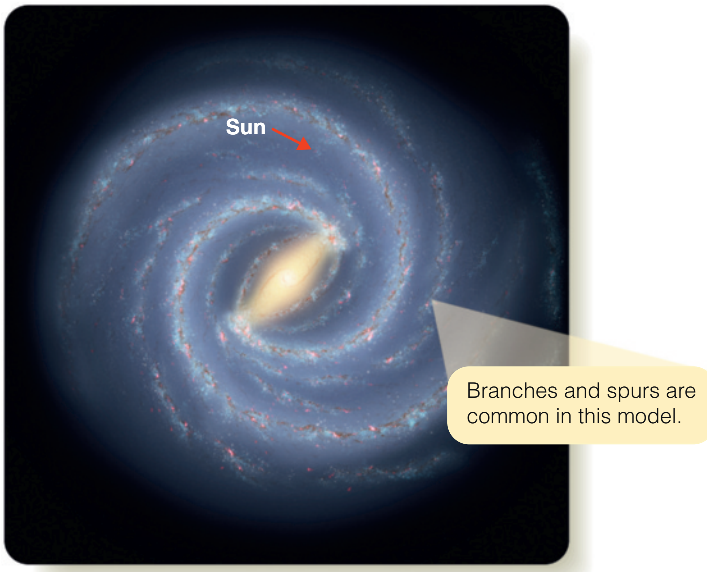


Figure 17-15

This artist's impression of a two-armed model for the Milky Way is based on observations with the Spitzer Infrared Space Telescope. Notice the large central bar.

clouds that distort the radial velocities. Observations toward the center of the galaxy have an additional problem. There, orbital motions of gas clouds are perpendicular to the line of sight, and all of the radial velocities are zero. That is why the map in Figure 17-14 is empty in the wedge-shaped region directly toward the center.

Clearly we live in a spiral galaxy, but the spiral pattern appears to be slightly irregular, with many branches and gaps. The stars you see in Orion, for example, appear to be in a detached segment of a spiral arm, referred to as a "spur." By combining observations of our own galaxy with studies of other galaxies, astronomers can make educated guesses about what the Milky Way Galaxy might look like viewed from outside (Figure 17-15). In addition to helping map the arrangement of spiral arms, infrared observations indicate that the central bulge of our galaxy is not a sphere but rather an elongated bar pointing partially away from Earth. In the next chapter, you will see that such structures are common in other galaxies.

One important fact revealed by the radio maps is that spiral arms are regions of higher gas density. Spiral tracers tell you that the arms contain young objects, so you suspect they must also have active star formation. Radio maps confirm your suspicion by telling you that the material needed to make stars is abundant and concentrated in spiral arms.

The Spiral Density Wave Theory

Having mapped the spiral pattern of our galaxy, and seen these patterns in other galaxies, you might ask, "Just what are spiral arms?" You can be sure they are not physically connected structures like bands of magnetic field that hold the gas in place. If they were, the differential rotation of a galaxy would destroy them within just a few hundred million years, winding them up and tearing them apart like paper streamers caught on the wheel of a speeding car. Yet spiral arms are common in galaxies (Figure 17-12), so they must last for billions of years.

Astronomers conclude that spiral arms are dynamically stable—they retain the same appearance even though the gas, dust, and stars in them are constantly changing. To see how this works, think of the traffic jam behind a slow-moving truck. Seen from an airplane, the traffic jam would be stable, moving slowly down the highway. But you could watch an individual car approach from behind, slow down, wait its turn, finally reach the front of the jam, pass the truck, and resume speed. The individual cars in the jam are constantly changing, but the traffic jam itself is a stable pattern, moving at its own speed along the highway.

In the **spiral density wave theory,** spiral arms are dynamically stable regions of compression that move slowly around the galaxy, just as the truck moves slowly down the highway. Gas clouds moving at orbital velocity around the galaxy overtake the slow-moving arms from behind and slam into the gas already in the arms. The sudden compression of the gas can trigger the collapse of the gas clouds and the formation of new stars (look back again to Chapter 14). Newly formed stars and the remaining gas eventually move on through the arm and emerge from the front of the slow-moving arm to resume their travels around the galaxy (Figure 17-16).

What evidence do astronomers have that star formation is caused by spiral density waves? Spiral tracers are the key. Stars of all masses are forming in spiral arms, but, as you have already learned, the O and B stars live such short lifetimes that they die before they can move out of the spiral arm. These massive, high-luminosity stars, along with gas and dust, confirm that spiral arms are sites of star formation.

You can expect that lower-mass stars like the sun can't be used as spiral tracers. They also form in spiral arms, but, during their relatively long lives, they leave the arms in which they were born. The sun probably formed as part of an association in a star-forming region within a spiral arm almost 5 billion years ago, escaped from that association, and has circled the galaxy more than 20 times, passing through many spiral arms. Plotting the locations of sunlike stars does not reveal the spiral pattern, confirming this picture.

The evidence seems to fit the spiral density wave theory well, but the theory has two problems. First, how does the complicated spiral disturbance begin, and how is it sustained? Computer models indicate that spiral density waves should slowly die out in about a billion years, so something must regenerate the spiral wave. Mathematical models show that the disk of the galaxy is naturally strongly affected by certain types of disturbances. As you learned earlier, observations show that the center of our galaxy is not a sphere but a bar. The gravitational effects of the bar's rotation could continuously perturb the galaxy's disk and

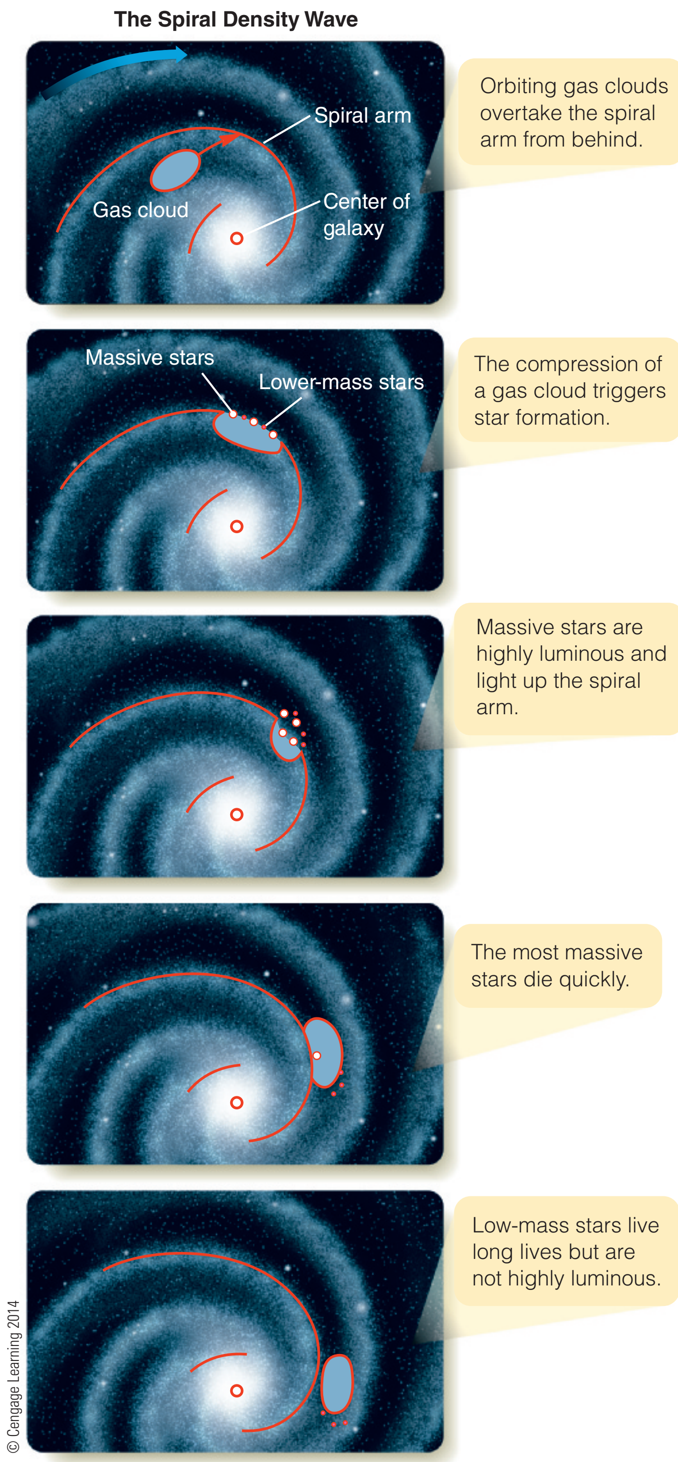


Figure 17-16

According to the spiral density wave theory, star formation occurs as gas clouds pass through spiral arms.

stimulate the formation of spiral density waves. A close encounter with a passing galaxy could also stimulate a spiral pattern. Some galaxies without a visible bar or large companion nevertheless have prominent spiral arms, so astronomers continue to make observations and computer models, investigating the question of spiral arm formation.

The second problem for the density wave theory involves the spurs and branches observed in the arms of our own and other galaxies. Some galaxies, called **grand-design** galaxies, have symmetric two-armed patterns (■ Figure 17-17a). Other galaxies have a great many short spiral segments, giving them a fluffy appearance, but no overall grand design. These galaxies have been termed **flocculent,** meaning "woolly" (Figure 17-17b). Our Milky Way Galaxy seems to be intermediate between the extremes of flocculent versus grand-design. What process might make small-scale spiral segments instead of a galaxy-wide spiral pattern? The answer may lie in self-sustaining ("contagious") star formation. You learned in Chapter 14 that star formation in one location can lead, by supernova explosions, stellar winds, and outflows, to star formation in neighboring locations (look back especially to Figure 14-9). The differential rotation of the galaxy will drag the inner edge of a star-forming cloud ahead and let the outer edge lag behind. This will result in a region of star formation shaped like a segment of a spiral arm: a spur (■ Figure 17-18). Astronomers suspect that while a spiral density wave can generate beautiful two-armed patterns, the **self-sustaining star formation** process produces the branches and spurs so prominent in some galaxies, including our own.

This discussion of star formation in spiral arms illustrates the importance of natural processes. The spiral density wave creates graceful arms, but it is the star formation in the arms that makes them stand out so prominently. Self-sustaining star formation can act in some galaxies to modify the spiral arms and produce branches and spurs. In other galaxies, it can make the spiral pattern flocculent. By searching out and understanding the details of such natural processes, astronomers can begin to understand the overall structure and evolution of the universe we live in. (**How Do We Know? 17-2**).

Figure 17-17

(a) Some galaxies are dominated by two spiral arms, but, even in these galaxies, minor spurs and branches are common. Spiral density waves can generate the two-armed, grand-design pattern, but self-sustained star formation may be responsible for the irregularities. (b) Many spiral galaxies do not appear to have prominent spiral arms. Abundant spurs and branches nevertheless suggest that star formation is proceeding robustly in such galaxies. Observations indicate that our Milky Way Galaxy's spiral pattern is intermediate between these two examples.

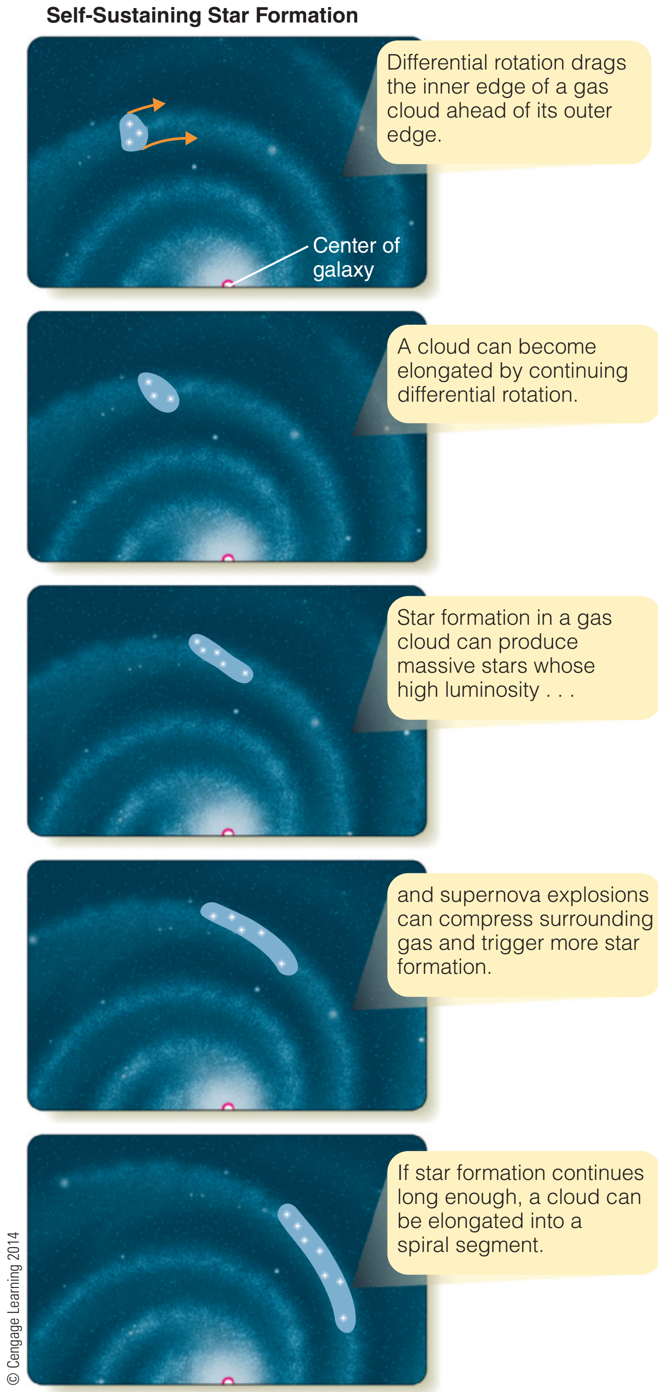


Figure 17-18

Self-sustaining star formation may be able to produce long clouds of young stars that look like segments of spiral arms.

SCIENTIFIC ARGUMENT

Why can't astronomers use solar-type stars as spiral tracers?
Sometimes the timing of events is the critical factor in a scientific argument. In this case, you need to think about the evolution of stars and their orbital periods around the galaxy. Stars like the sun live about 10 billion years, but the sun's orbital period around the galaxy is 210 million years. The sun almost certainly formed when a gas cloud passed through a spiral arm, but since then the sun has circled the galaxy many times and has passed through spiral arms often. That means the sun's present location, which happens coincidentally to be near a spiral arm, is not due to the location of that arm. An O star, however, lives only a few million years. It is born in a spiral arm and lives out its entire lifetime before it can leave the spiral arm. Short-lived stars such as O stars are found only in spiral arms, but G stars are found all around the galaxy.

The spiral arms of our galaxy would make it beautiful if it could be photographed from a distance, but we are trapped inside it. Create an argument based on evidence: **How do astronomers know that the spiral arms mapped out near us by spiral tracers actually extend across the disk of our galaxy?**

17-4 The Nucleus of the Galaxy

THE MOST MYSTERIOUS REGION of our galaxy is its very center, the nucleus. At visual wavelengths, this region is totally hidden by dust and gas that dim the light it emits by 30 magnitudes. If a trillion (10^{12}) photons of light left the center of the galaxy on a journey to Earth, only one of those photons would make it through the dust and gas. Consequently, visual-wavelength images reveal nothing about the nucleus. Observations at infrared and radio wavelengths can see through the interstellar material, and those images show a region of tremendously crowded stars orbiting in the nucleus at high velocity. To understand what is happening in the innermost regions of our galaxy, you need to carefully compare observations and hypotheses.

Observations of the Galactic Nucleus

If you look up at the Milky Way on a dark night, you might notice a slight thickening in the direction of the constellation Sagittarius, but nothing specifically identifies this as the direction of the heart of the galaxy. Even Shapley's study of globular clusters identified the location only approximately. The first complete infrared map of the central bulge made by Eric Becklin in 1968 showed the location of intense radiation where the stars are most crowded together, the gravitational core of the galaxy. High-resolution radio maps of that stellar center revealed a complex collection of radio sources, with one, **Sagittarius A***

How Do We Know? 17-2

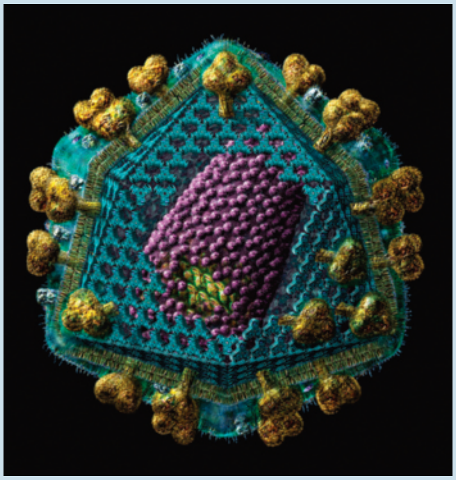
Nature as Processes

How is getting a cold like stars building the chemical elements? Science, at first glance, seems to be nothing but facts, but in many cases, you can organize the facts into the story of a process. For example, astronomers try to assemble the sequence of events that led to the formation of the chemical elements. If you understand that process, you have command over a lot of important facts in astronomy.

A process is a sequence of events that leads to some result or condition, and much of science is focused on understanding how these natural processes work. Biologists, for example, try to understand how a virus reproduces. They must figure out how the virus tricks the immune system into leaving it alone, penetrates the wall of a healthy cell, injects viral DNA, commandeers the cell's resources to make new viruses, and finally destroys the cell to release the new virus copies. A biologist may spend her life studying a specific step, but the ultimate goal of science is to tell the entire story of the process.

As you study any science, be alert for processes as organizing themes. When you see a process in science, ask yourself a few basic questions. What conditions prevailed at the beginning of the process? What sequence of steps occurred? Can some steps occur simultaneously, or must one step occur before another can occur? What is the final state that this process produces?

Recognizing a processes and learning to tell its story will help you remember a lot of details, but that is not the real value of a scientific process. Identifying a process and learning to tell its story helps you understand how nature works and explains why the universe is the way it is.



Russell Knightly Media, km.com.au

A virus is a collection of molecules that cannot reproduce until it penetrates into a living cell.

(abbreviated **Sgr A*** and usually pronounced "sadge A-star"), lying at the exact location of the galactic nucleus.

Radio interferometer observations show that Sgr A* is less than 1 AU in diameter but nevertheless is a strong source of radio energy as well as X-ray emission (short-wavelength X-rays can also penetrate interstellar dust and gas). Look back to the opening illustration for this chapter. There you will see an image of the central 70 pc of the galaxy that combines near-infrared (represented by yellow), far-infrared (red) and X-ray (blue and green) data from three space telescopes. You can see that powerful forces are at work in the nuclear region. The tremendous amount of infrared radiation coming from the central area appears to be produced by crowded stars and by dust warmed by those stars. But what could be as small as Sgr A* and produce so much radio and X-ray energy?

Read **Sagittarius A*** on pages 386–387 and notice three important points:

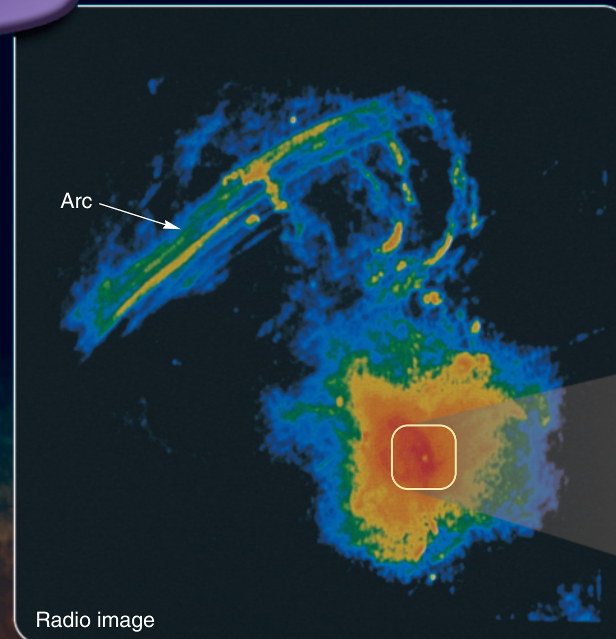
1. Observations at radio and infrared wavelengths reveal complex structures near Sgr A* caused by magnetic fields and by rapid star formation. Supernova remnants show that massive stars have formed there recently and exploded at the ends of their lives.
2. The center is crowded. Tremendous numbers of stars heat the dust, which emits strong infrared radiation.
3. Finally, there is evidence that Sgr A* is a supermassive black hole into which gas is flowing. Observations of the motions of stars orbiting the central object indicate its mass is at least 4 million $M_\odot$.

A supermassive black hole is an exciting idea, but scientists must always be aware of the difference between adequacy and necessity. A supermassive black hole is adequate to explain the observations, but is it necessary? Could there be other explanations? For example, some astronomers have suggested that gas flowing inward could trigger tremendous bursts of star formation, causing some of the turmoil observed in the galactic nucleus. Such hypotheses have been considered and tested against the evidence, but none appears to be adequate to explain all the observations. So far the only viable hypothesis is that our galaxy's nucleus is home to a supermassive black hole.

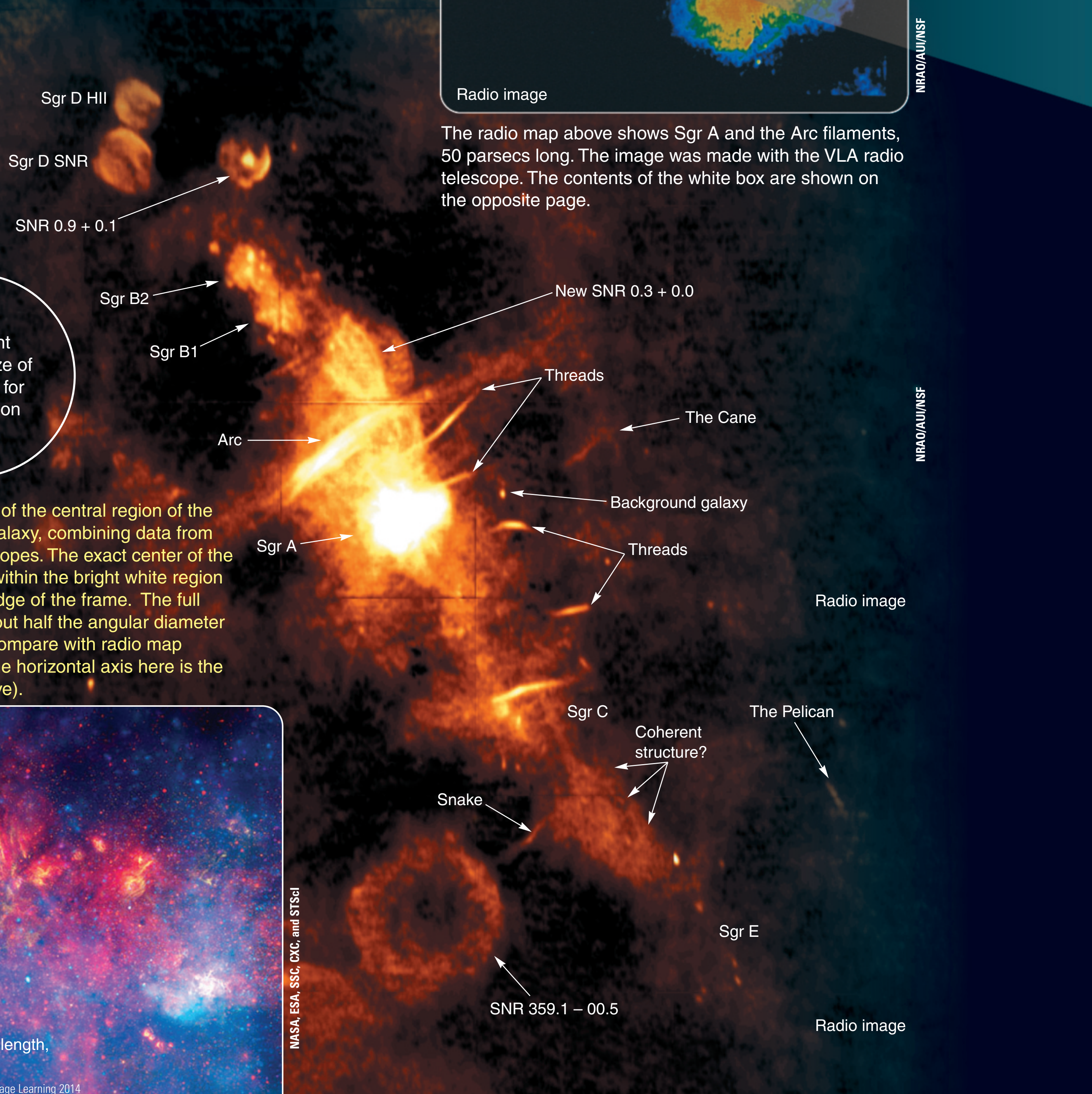
Meanwhile, observations are allowing astronomers to refine their models. For instance, Sgr A* is not as bright in X-rays as it should be if it has a hot accretion disk with matter constantly flowing into the black hole. Observations of X-ray and infrared flares lasting only a few hours suggest that mountain-size blobs of matter may occasionally fall into the black hole and be ripped apart and heated by tides. But the black hole seems to be mostly dormant, lacking a fully developed hot accretion disk because little matter is flowing into it at the present time.

Sagittarius A*

1 There is so much interstellar dust in the plane of the Milky Way that you cannot observe the nucleus of our galaxy at visual wavelengths. The image below is a radio image of the innermost 300 pc. Many of the features are supernova remnants (labeled SNR), and a few are star formation clouds. Peculiar features such as threads, the Arc, and the Snake may be gas trapped in magnetic fields. At the center of the image lies Sagittarius A (Sgr A), the location of the nucleus of our galaxy.



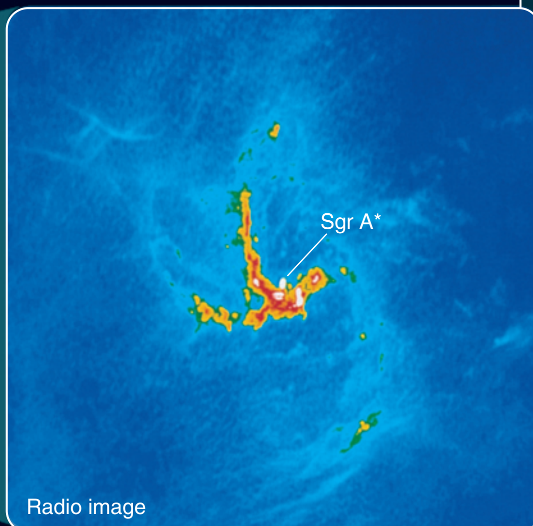
The radio map above shows Sgr A and the Arc filaments, 50 parsecs long. The image was made with the VLA radio telescope. The contents of the white box are shown on the opposite page.



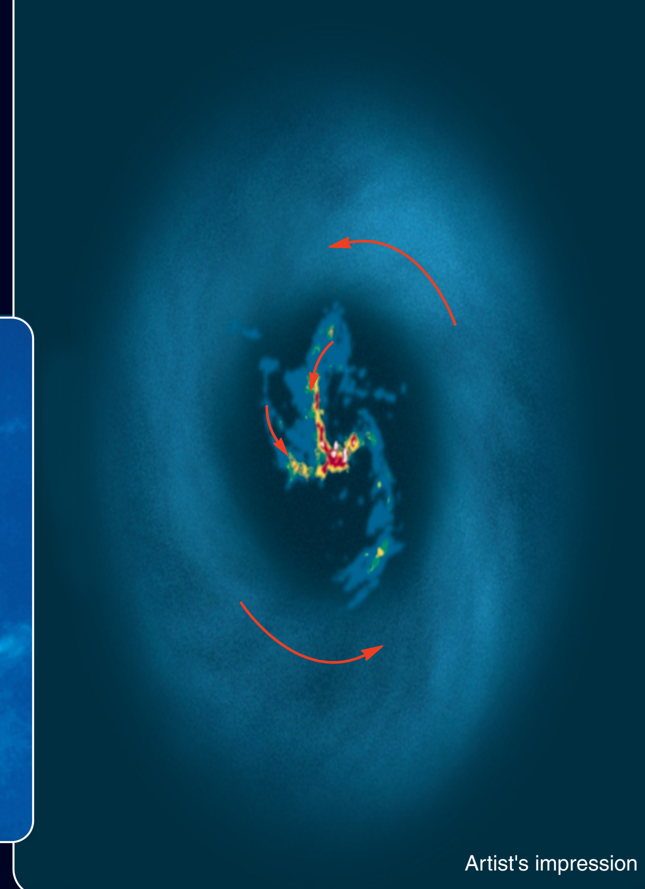
2 An image of the central region of the Milky Way Galaxy, combining data from three space telescopes. The exact center of the galaxy is located within the bright white region to near the right edge of the frame. The full image width is about half the angular diameter of the full moon (compare with radio map above; note that the horizontal axis here is the diagonal axis above).

Infrared, visual-wavelength, and X-ray image

1a This high-resolution radio image of Sgr A (within the white box in the small-scale map on the opposite page) reveals a spiral swirl of gas around an intense radio source known as Sgr A*, about 3 pc across, this spiral lies in a low-density cavity inside a larger disk of neutral gas. Astronomers have evidence that this small spiral is composed of streams of matter flowing into Sgr A* from the inner edge of the larger disk (drawing at right).



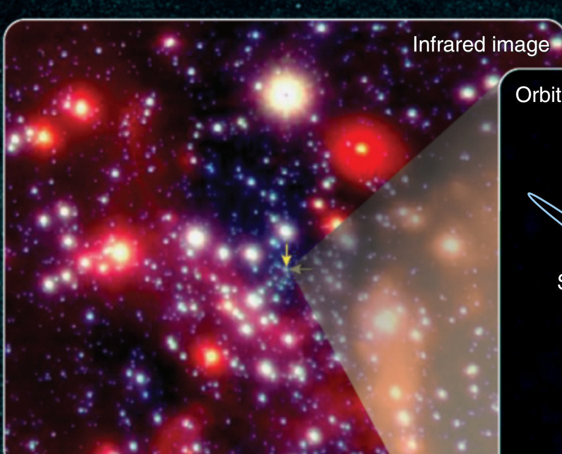
N. Killeen and Kwok-Yung Lo/NRAO



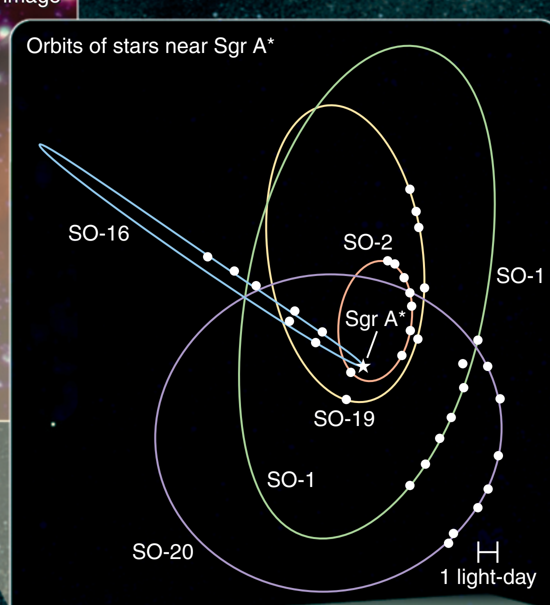
N. Killeen and Kwok-Yung Lo/NRAO

Evidence of a Black Hole in the Nucleus of Our Galaxy

3 Since the middle 1990s, astronomers have been able to use large infrared telescopes and adaptive optics to follow the motions of stars orbiting around Sgr A*. A few of those orbits are shown here. The size and period of the orbit allows astronomers to calculate the mass of Sgr A* using Kepler's third law. The orbital period of the star SO-2, for example, is 15.2 years, and the semimajor axis of its orbit is 950 AU. The combined motions of the observed stars suggest that Sgr A* has a mass of 4 million solar masses.



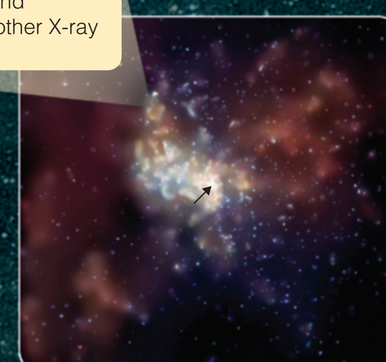
ESO/S. Gillessen et al.



At its closest, SO-2 comes within 17 light-hours of Sgr A*. The available evidence eliminates alternate hypotheses that Sgr A* is a large cluster of normal stars or neutron stars or stellar-mass black holes. Only a single black hole could contain so much mass in so small a region.

For comparison, the diameter of the planetary region of our solar system, defined by Neptune's orbit, is half a light-day.

The *Chandra X-Ray Observatory* has imaged Sgr A* and detected over 2000 other X-ray sources in the area.


NASA/CXC/MIT/F.K. Baganoff et al.

3a A black hole with a mass of 4 million solar masses would have an event horizon with a size on the scale of this diagram smaller than the period at the end of this sentence. A slow dribble of only 0.0002 solar mass of gas per year flowing into the black hole could produce the observed energy. A sudden increase, such as when a star falls in, could produce a violent eruption.

The evidence of a massive black hole at the center of our galaxy seems conclusive. It is much too massive to be the remains of a dead star, however, and astronomers conclude that it probably formed as the galaxy first took shape.

Such a supermasssive black hole could not be the remains of a single dead star. It contains too much mass. It probably formed when the galaxy first formed over 13 billion years ago. In later chapters you will see that such supermassive black holes are found at the centers of most large galaxies.

SCIENTIFIC ARGUMENT

Why do astronomers think the center of our galaxy contains a large mass?

Because scientific arguments usually hinge on evidence, they often involve discussions of measurement—one of the keys to science. The best way to measure the mass of an astronomical object is to watch something orbit around it. Then, you can use Kepler's third law to find the mass inside the orbit. Because of interstellar dust, astronomers can't see to the center of our galaxy at visual wavelengths, but infrared observations can detect individual stars orbiting Sgr A*. The star S0-2 has been particularly well observed, but a number of stars can be followed as they orbit the center. The sizes and periods of these orbits, interpreted using Kepler's laws, reveal that Sgr A* contains at least 4 million solar masses.

Now build a new argument to analyze a different observation. **The strong infrared radiation at wavelengths longer than 4 microns (4000 nm) implies vast numbers of stars are crowded into the center. Why?**

17-5 Origin and History of the Milky Way Galaxy

Just as dinosaurs left behind fossilized footprints, our galaxy has left behind clues about its youth. Astronomers have compelling evidence that the stars of the spherical component are old and must have formed long ago when the galaxy was very young. Some of that evidence comes from comparing abundances of chemical elements in stars located in different parts of the galaxy.

The Element-Building Process

You can recall what you have learned about stellar structure and stellar evolution in previous chapters to understand the chemical evolution of the galaxy. In astronomy jargon, the term **metals** refers to all of the chemical elements heavier than helium. (Note that this is not what the word *metals* means to nonastronomers.) In a later chapter you will learn about evidence that when the universe began it contained about 90 percent hydrogen atoms, 10 percent helium atoms, and little or no metals. The first stars that formed early in the universe's history therefore had to be nearly pure hydrogen and helium. All of the other chemical elements have been produced by **nucleosynthesis,** the process in stars that fuses hydrogen and helium to make the heavier atoms.

As you already know, medium-mass stars like the sun cannot ignite carbon fusion, but during helium fusion the heat and density can trigger some nuclear reactions that cook the gas to produce small amounts of elements heavier than helium. When the aging star pushes away its surface layers to produce a planetary nebula, some of those elements are spread back into the interstellar medium where they may become part of newly forming stars and planetary systems.

The most massive stars fuse elements up to iron and simultaneously cook the gas in their cores to produce small amounts of many different atoms, including sulfur and calcium. When those stars die in supernova explosions, traces of those elements are spread back into the interstellar medium, along with rarer atoms such as gold, platinum, and uranium produced in the supernova explosion itself, which can also be incorporated eventually into new stars and planets.

Nevertheless, metals are actually quite rare in the universe. Conventionally, astronomers graph the abundance of the elements using an exponential scale, but if you replot the data using a linear scale you can see how rare these atoms are (■ Figure 17-19).

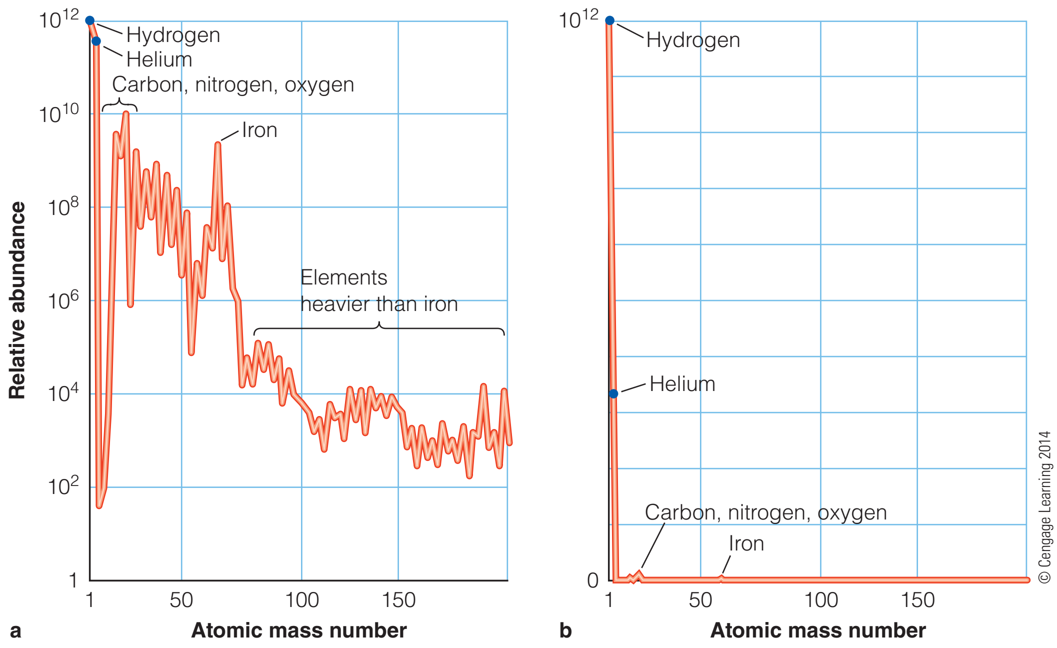
Stellar Populations

In the 1940s, astronomers realized that there are two families of stars in the galaxy. They form and evolve in similar ways, but they differ, especially in their abundances of metals. **Population I stars** are metal rich, containing 1 to 3 percent metals, whereas **Population II stars** are metal poor, typically containing less than 0.1 percent metals. The difference is clearly evident in spectra (■ Figure 17-20). The abundances plotted in Figure 17-19 show metal-rich Population I composition. H–R diagrams of star clusters (look back to Chapter 15) reveal that Population I stars are relatively young, and Population II stars are old.

Population I stars belong to the disk component of the galaxy and are sometimes called disk population stars. They have nearly circular orbits in the plane of the galaxy and formed within the last few billion years. The sun is a Population I star, as are the type I Cepheid variables discussed in this chapter.

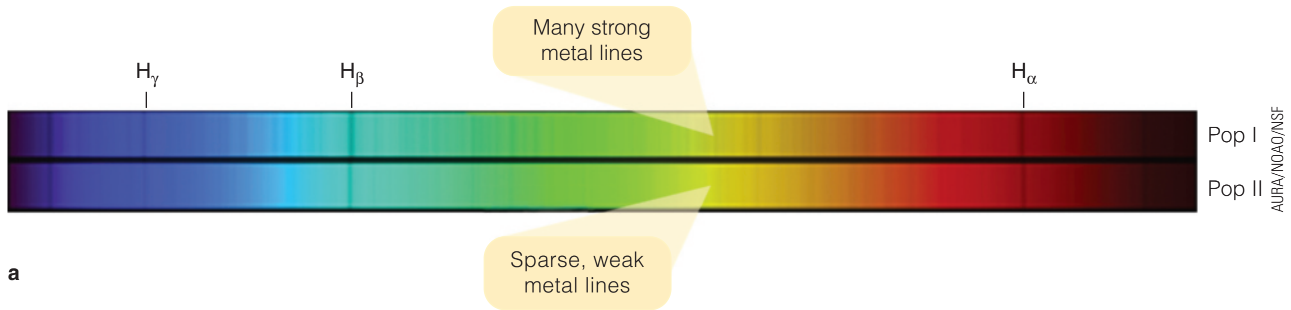
Population II stars belong to the spherical component of the galaxy and are sometimes called the halo population stars. These stars have randomly tipped orbits ranging from circular to highly elliptical. They are old stars that formed when the galaxy was young. The metal-poor globular clusters are part of the halo population, as are the RR Lyrae and type II Cepheids.

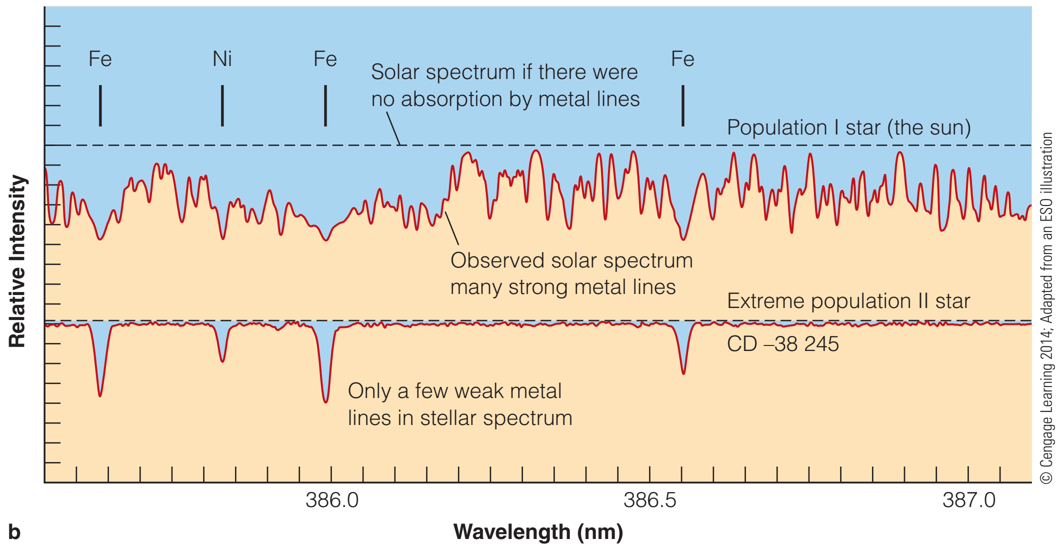
Further observations show that there is a gradation between populations. Extreme Population I stars are found only in the spiral arms. Slightly less metal-rich Population I stars, called intermediate Population I stars, are located throughout the disk. The sun is an intermediate Population I star. Stars that are even less metal rich, such as stars in the central bulge, belong to the intermediate Population II. The most metal-poor stars are those in the halo, including those in globular clusters. Those are referred to as extreme Population II stars.



■ **Figure 17-19**

The abundances of chemical elements in the universe. (a) When the elemental abundances are plotted on an exponential scale, you see that elements heavier than iron are about a million times less common than iron and that all elements heavier than helium (the metals) are quite rare. (b) The same data plotted on a linear scale provide a more realistic impression of how rare the metals are. Carbon, nitrogen, and oxygen make small peaks near atomic mass 15, and iron is just visible in the graph.





■ **Figure 17-20**

(a) The difference between spectra of Population I stars and Population II stars is dramatic. Examine the upper spectrum here and notice the hundreds of faint spectral lines. The lower spectrum has fewer and weaker lines. (b) A graph of such spectra reveals overlapping absorption lines of metals completely blanketing the Population I spectrum. The lower spectrum is that of an extremely metal-poor star with only a few weak metal lines of iron (Fe) and nickel (Ni). This Population II star contains about 10,000 times less metal than the sun.

■ **Table 17-1** | **Stellar Populations**

	Population I		Population II	
	Extreme	Intermediate	Intermediate	Extreme
Location	Spiral arms	Disk	Central bulge	Halo
Metals (%)	3	1–2	0.1–1	Less than 0.1
Shape of orbit	Circular	Slightly eccentric	Moderately eccentric	Highly eccentric
Average age (yr)	0.2 billion and younger	0.2–5 billion	5–10 billion	10–13 billion

The two populations of stars in our galaxy are clearly different (■ Table 17-1). Those differences are clues to the history of the galaxy.

When you observe Population II stars, such as those in the halo, you are looking at the survivors of the early generations of stars in our galaxy. The first stars formed from gas that was metal poor, and the only survivors of these early generations are low-mass, long-lived stars. Their spectra, which indicate the composition of their atmospheres and not their cores, still show the composition of the metal-poor gas from which they formed. Population I stars such as the sun formed more recently, after the interstellar medium had been enriched in metals, and their spectra show stronger metal lines. Stars forming now have even higher metal abundances.

Age of the Galaxy

Because astronomers know how to find the age of star clusters, they can estimate the age of the oldest stars in the galaxy, giving a lower limit to the age of the entire galaxy. The process sounds straightforward, but uncertainties make the easy answer hard to interpret.

The oldest open clusters are 9 to 10 billion years old. These ages come from the turnoff points in their H–R diagrams, but finding the age of an old cluster is difficult because old clusters change so slowly. Also, the exact location of the turnoff point depends on chemical composition, which differs slightly among clusters. Finally, open clusters are not strongly bound by gravity, so older open clusters may have dissipated as their stars wandered away. It is reasonable to suppose that the galactic disk is somewhat older than the oldest remaining open clusters, which suggests that the disk is at least 10 billion years old.

Globular clusters have faint turnoff points in their H–R diagrams and are clearly old, but finding these ages is difficult. Again, slight differences in chemical composition make noticeable differences when calculating the stellar models from which ages are determined. Also, to find the age of a cluster, astronomers must know its distance. Parallaxes from the *Hipparcos* satellite may allow astronomers to improve the calibration of the Cepheid and RR Lyrae variable stars, and careful studies with the newest large telescopes have refined the chemical compositions and better defined the H–R diagrams of globular clusters. The best analysis of all of the data suggests that the average globular clusters are about 11 billion years old, although some globular clusters are younger than that, and some are older. Studies of the oldest globular clusters suggest that the halo of our galaxy is at least 13 billion years old.

Observations of stellar populations and clusters show that the disk is younger than the halo. You can combine these ages with the process of nucleosynthesis to tell the story of our galaxy.

Formation of the Galaxy

The lack of metals in the spherical component of the galaxy tells you it is very old, a fossil left behind by the galaxy when it was young and drastically different from its present disk shape. The study of element building and stellar populations leads to the fundamental question, "How did our galaxy form?"

In the 1950s, astronomers began to develop a hypothesis sometimes called the **monolithic collapse** hypothesis, also known as the "**top-down**" hypothesis, to explain the formation of our galaxy (■ Figure 17-21). Recent observations, however, are forcing a reevaluation of that hypothesis.

The monolithic collapse hypothesis says that the galaxy formed from a single large cloud of turbulent gas over 13 billion years ago. That cloud contracted to form our galaxy. As gravity pulled the gas inward, the cloud began to fragment into smaller clouds, and because the gas was turbulent, the smaller clouds had random velocities. That caused the stars and star clusters that formed from these fragments to have orbits with a wide range of shapes; a few were circular, most were eccentric, and some were extremely eccentric. The orbits were also inclined at different angles, resulting in a spherical cloud of stars—the spherical component of the galaxy. Of course, these first stars were metal poor because no stars had existed earlier to enrich the gas with metals.

The second stage in this hypothesis accounts for the formation of the disk component. As the gas clouds in the originally spherical cloud collided, turbulent motions in the gas canceled out, as do eddies in recently stirred coffee, and the cloud was left with a uniform rotation. A rotating, low-density cloud of gas

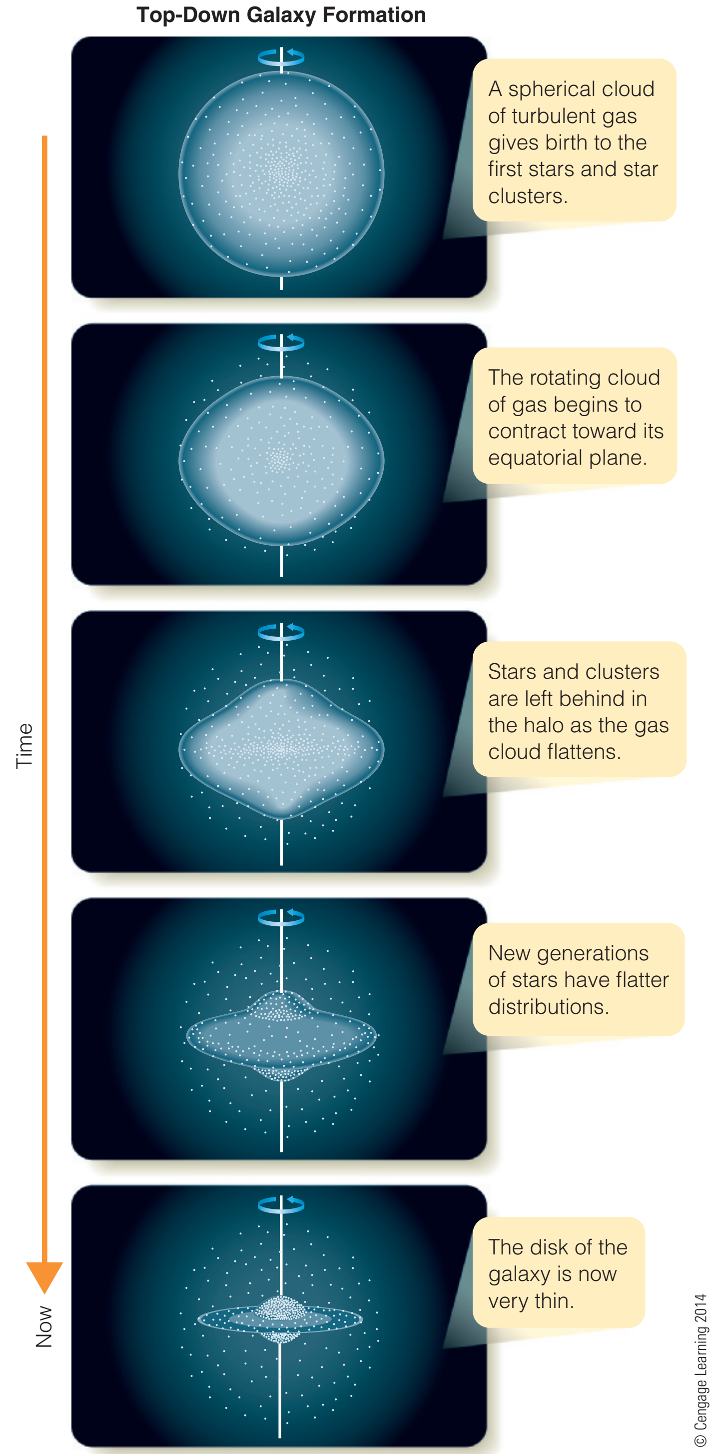


Figure 17-21

The "monolithic collapse", or top-down, model for the origin of our galaxy begins with a spherical gas cloud that flattens into a disk.

cannot remain spherical. A star is spherical because its high internal pressure balances its gravity, but in a low-density cloud, the internal pressure is much too low to support the weight. Like a blob of pizza dough spun in the air, the cloud must flatten into a disk (Figure 17-21), eventually producing the galaxy's disk.

This contraction from a sphere into a disk would have taken billions of years, and while that happened the metal abundance gradually increased as generations of stars were born from the gradually flattening gas cloud. The stars and globular clusters of the halo were left behind by the cloud as it flattened, and subsequent generations of stars formed in flatter distributions. The gas distribution in the galaxy now is so flat that the youngest stars are confined to a disk only about 100 parsecs thick. These stars are metal rich and have nearly circular orbits.

The monolithic collapse hypothesis accounts for many of the Milky Way's properties. Advances in technology, however, have improved astronomical observation, and beginning in the 1980s, contradictions between theory and observation arose. For example, not all globular clusters have the same age, but surprisingly, some of the youngest clusters seem to be in the outer halo. In contrast, the monolithic collapse hypothesis says that the halo formed first and predicts that the clusters within it should either have a uniform age or the most distant ones should be somewhat older. Another problem is that the oldest stars are observed to be metal poor but not completely metal free. There must have been at least a few massive stars to create these metals during a generation before the formation of the oldest stars now seen in the halo.

Can the original hypothesis be modified to explain these observations? Astronomers suspect that the galaxy began as a large gas cloud containing almost no metals. Models of metal-free stars show that they are very massive, so that first generation of stars evolved rapidly and exploded as supernovae, which enriched the gas cloud with traces of metals. None of those massive first-generation stars survive, but the metals they created are detectable in the oldest Population II stars.

It seems likely that the central bulge and halo formed from such a gas cloud or from the accumulation of a number of gas clouds. A thick but low-density disk of stars may have formed at this early stage. This would explain the age of the central bulge and the metals in both the oldest stars in the halo and those scattered above and below the disk. The thin part of the disk could have formed later as more gas fell into the galaxy and settled into the gas clouds of the thin disk where stars are forming today. Perhaps entire galaxies were captured by the growing Milky Way. Astronomers have found streams and rings of scattered stars surrounding our galaxy and hypothesize that they were produced when smaller galaxies were captured, pulled apart, and absorbed by our home galaxy. You will see dramatic evidence in the next chapter that such galaxy mergers do occur.

If the young Milky Way Galaxy in fact was partly assembled from smaller units, absorbing several small but partially evolved galaxies, then some of the globular clusters astronomers see in the halo may be hitchhikers. Astronomers now have evidence that our galaxy formed through merger of a number of smaller pre-galaxy clouds of gas and stars, plus later additions of infalling gas and more small galaxies (■ Figure 17-22). This **bottom-up hypothesis** for the formation of the galaxy would explain puzzles

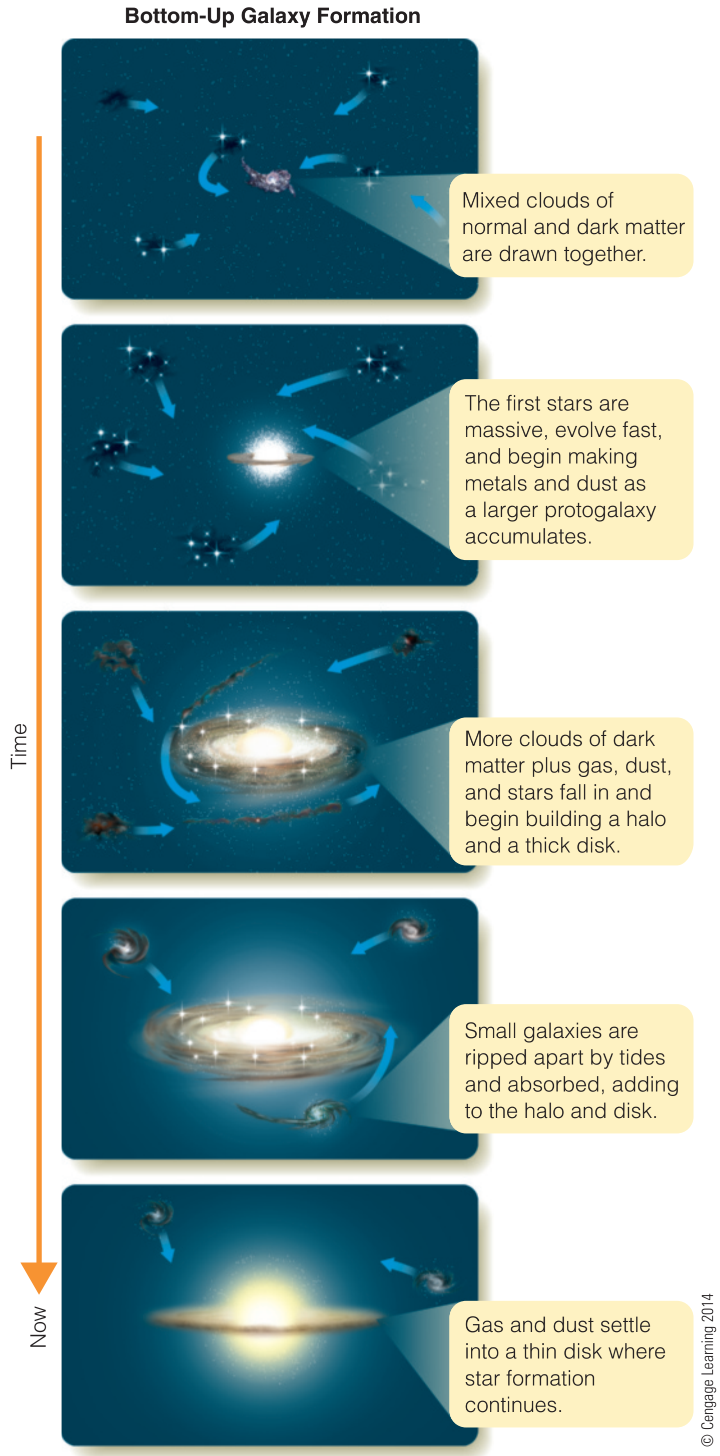


Figure 17-22

The bottom-up hypothesis for the formation of the Milky Way Galaxy proposes that smaller star systems accumulated to form larger ones. To see how this could have built our galaxy, start with frame (a), only a few hundred million years after the beginning of the universe, as small clouds of matter begin accumulating and stars begin forming in them. In frame (b) the central object has grown larger, and in frame (c) the galactic halo and disk are forming. By frame (e), representing today, the disk of the galaxy has become very thin.

such as the observed ranges of globular cluster ages and metallicities. The metal abundances and ages of stars and star clusters in our galaxy seem to be important clues, but metal abundance and age do not tell the whole story. Astronomer Bernard Pagel was thinking of this when he said, "Cats and dogs may have the same age and metallicity, but they are still cats and dogs."

SCIENTIFIC ARGUMENT

Why do metal-poor stars have the most eccentric orbits?

A good argument makes connections between ideas, and sometimes those connections are not obvious. Certainly, the metal abundance of a star cannot affect its orbit, so an analysis must not confuse cause and effect with the relationship between these two factors. Both chemical composition and orbital shape depend on a third factor—age. The oldest stars are metal poor because they formed before there had been many supernova explosions to create and scatter metals into the interstellar medium. Those stars formed long ago when the galaxy was young and motions were not organized into a disk, so the stars tended to take up randomly shaped orbits, many of which are quite elongated and eccentric. Consequently, today, the most metal-poor stars tend to follow the most eccentric orbits.

Nevertheless, even the oldest stars known in our galaxy contain some metals. They are metal poor, not metal free. Adjust your argument. **Where did these metal-poor stars get their metals?**

What Are We?
Children of the Milky Way

Hang on tight. The sun, with Earth in its clutch, is ripping along at about 240 km/sec (that's 540,000 mph) as it orbits the center of the Milky Way Galaxy. We live on a wildly moving ball of rock in a large galaxy that humanity calls home, but the Milky Way is more than just our home. Perhaps "parent galaxy" would be a better name.

Except for hydrogen atoms, which have survived unchanged since the universe began, you and Earth are made of metals—atoms heavier than helium. There is no helium in your body, but there is plenty of carbon, nitrogen, and oxygen. There is calcium in your bones and iron in your blood. All of those atoms and more were cooked up inside stars or in their supernova deaths.

Stars are born when clouds of gas orbiting the center of the galaxy collide with the gas in spiral arms and are compressed. That process has given birth to generations of stars, and each generation has produced elements heavier than helium and spread them back into the interstellar medium. The abundance of metals has grown slowly in the galaxy. About 4.6 billion years ago a cloud of gas enriched in those heavy atoms slammed into a spiral arm and produced the sun, Earth, and you. You have been created by the Milky Way—your parent galaxy.

Study and Review

Summary

- The hazy band of the Milky Way is our wheel-shaped galaxy seen from within, but its size and shape are not obvious. William and Caroline Herschel counted stars at many locations over the sky to show that our star system seemed to be shaped like a grindstone with the sun near the center.
- Later astronomers studied the distributions of stars, but because they didn't know that interstellar dust blocked their view of distant stars, they incorrectly concluded the star system is only about 10 **kiloparsecs (kpc) (p. 369)** in diameter with the sun at the center.
- In the early 20th century, Harlow Shapley observed variable stars called **Cepheids (p. 370)** and **RR Lyrae stars (p. 371)** in an effort to determine the size of the galaxy. Cepheids and RR Lyraes fall within an **instability strip (p. 370)** in the H–R diagram. The **period–luminosity relation (p. 371)** for Cepheid variables originally discovered by Henrietta Leavitt was **calibrated (p. 372)** by Shapley using the **proper motions (p. 372)** of a few Cepheids. That allowed him to estimate the distance to globular clusters, demonstrating that our galaxy is much larger than the portion that can easily be observed from Earth and that the sun is not at the center.
- Modern observations suggest that our galaxy contains a **disk component (p. 375)** about 80,000 ly in diameter and that the sun is two thirds of the way from the center to the visible edge. The **central bulge (p. 376)** around the center and an extensive **halo (p. 376)** containing old stars and little gas and dust make up the **spherical component (p. 376).**
- The mass of the galaxy can be found from its **rotation curve (p. 377).** Kepler's third law reveals that the galaxy contains at least 110 billion solar masses. If stars orbited in **Keplerian motion (p. 377),** more distant stars would orbit more slowly. They do not, and that shows that the halo may contain much more mass than is visible. Because the mass in this **galactic corona (p. 377)** is not emitting detectable electromagnetic radiation, astronomers call it **dark matter (p. 378).**
- You can trace the spiral arms through the sun's neighborhood by using **spiral tracers (p. 379)** such as OB associations, but to extend the map over the entire galaxy, astronomers must use radio, infrared, and X-ray telescopes to see through the interstellar medium.
- The most massive stars live such short lives that they don't have time to move from their place of birth. Because they are found scattered along the spiral arms, astronomers conclude that the spiral arms are sites of star formation.
- The **spiral density wave theory (p. 381)** suggests that the spiral arms are regions of compression that move around the disk. When an orbiting gas cloud overtakes the compression wave, the gas cloud is compressed and forms stars. A density wave produces a two-armed **grand-design (p. 382)** spiral galaxy.
- Another process, **self-sustaining star formation (p. 382),** may act to modify the arms with branches and spurs as the birth of massive stars triggers the formation of more stars by compressing neighboring gas clouds. This may account for the ragged appearance of **flocculent (p. 382)** galaxies.
- The nucleus of the galaxy is invisible at visual wavelengths, but radio, infrared, and X-ray radiation can penetrate the gas and dust. These wavelengths reveal crowded central stars and warmed dust.
- The very center of the Milky Way Galaxy is marked by a radio source, **Sagittarius A* (Sgr A*) (p. 384).** That object must be less than 1 AU in diameter, but the motions of stars around the center show that it must contain at least approximately 4 million solar masses. A supermassive black hole is the only object that could contain so much mass in such a small space.
- The oldest star clusters reveal that the disk of our galaxy is younger than the halo, and the oldest globular clusters appear to be about 13 billion years old. So our galaxy must have formed about 13 billion years ago.
- Stellar populations are an important clue to the formation of our galaxy. The first stars to form, termed **Population II stars (p. 388),** were poor in elements heavier than helium—elements that astronomers call **metals (p. 388).** As generations of stars manufactured metals in a process called **nucleosynthesis (p. 388)** and spread them back into the interstellar medium, the metal abundance of more recent generations increased. **Population I stars (p. 388),** including the sun, are richer in metals.
- Because the halo is made up of Population II stars and the disk is made up of Population I stars, astronomers conclude that the halo formed first and the disk later.
- The **monolithic collapse,** or **top-down (p. 390)** hypothesis—that the galaxy formed from a single, roughly spherical cloud of gas and gradually flatted into a disk—does not match all of the current evidence. A newer **"bottom-up" hypothesis (p. 391)** includes mergers with smaller galaxies plus infalling gas continuing to contribute to the disk.

Review Questions

1. Why is it difficult to specify the dimensions of the disk and halo?
2. Why didn't astronomers before Shapley realize how large the galaxy is?
3. What evidence can you cite that our galaxy has a galactic corona?
4. Contrast the motion of the disk stars and that of the halo stars. Why do their orbits differ?
5. Why are all spiral tracers young?
6. Why couldn't spiral arms be physically connected structures? What would happen to them?
7. What are some hypotheses for mechanisms to trigger or sustain spiral density waves?
8. Why does self-sustaining star formation produce clouds of stars that look like segments of spiral arms?
9. Describe the kinds of observations you would make to study the galactic nucleus.
10. Why must astronomers use infrared telescopes to observe the motions of stars around Sgr A*?
11. What evidence can you cite that the nucleus of the galaxy contains a supermassive black hole?
12. Why are metals less abundant in older stars than in younger stars?
13. Why do metal-poor stars have a wider range of orbital shapes than metal-rich stars like the sun?
14. What evidence contradicts the monolithic collapse hypothesis for the origin of our galaxy?
15. **How Do We Know?** Calibration simplifies complex measurements, but how does that make the work of later astronomers dependent on the expertise of the astronomer who did the calibration?
16. **How Do We Know?** The story of a process makes the facts easier to remember, but that is not the true goal of the scientist. What is the real value of understanding a scientific process?

Discussion Questions

1. How would the information in this chapter differ if interstellar dust did not block starlight?
2. Why doesn't the Milky Way circle the sky along the celestial equator or the ecliptic?

Problems

1. Make a scale sketch of our galaxy in cross section. Include the disk, sun, nucleus, halo, and some globular clusters. Try to draw the globular clusters to scale.
2. Because of interstellar dust, astronomers can see at most about 5 kpc into the disk of the galaxy at visual wavelengths. What percentage of the galactic disk does that include? (*Hint:* Consider the area of the entire disk versus the area visible from Earth.)
3. If the fastest passenger aircraft can fly 0.45 km/s (1000 mph), how long would it take to reach the sun? The galactic center? (*Note:* 1 pc = 3.1×10^{13} km.)
4. If a typical halo star has an orbital velocity of 250 km/s, how long does it take to pass through the disk of the galaxy? Assume that the disk is 1000 pc thick.
5. If the RR Lyrae stars in a globular cluster have apparent magnitudes of 14, how far away is the cluster? (*Hints:* See Figure 17-4, and use the magnitude-distance formula in Chapter 13.)
6. If interstellar dust makes an RR Lyrae variable star look 1 magnitude fainter than it should, by how much will you overestimate its distance? (*Hint:* Use the magnitude-distance formula in Chapter 13.)
7. If you assume that a globular cluster 4.0 arc minutes in diameter is actually 25 pc in diameter, how far away is it? (*Hint:* Use the small-angle formula, Chapter 3.)
8. If the sun is 4.6 billion years old, how many times has it orbited the galaxy?
9. If astronomers were to find they have made a mistake and our solar system is actually 7.2 (rather than 8.2) kpc from the center of the galaxy, but the orbital velocity of the sun is still 240 km/s, what is the minimum mass of the galaxy? (*Hint:* Use Kepler's third law, Chapter 4.)
10. What temperature would interstellar dust have to have to radiate most strongly at 100 μm? (*Note:* 1 μm = 1000 nm. *Hint:* Use Wien's law, Chapter 6.)
11. Infrared radiation from the center of our galaxy with a wavelength of about 2 μm (2×10^{-6} m) comes mainly from cool stars. Use this wavelength as λ_{max} and find the temperature of the stars.
12. If an object at the center of the Milky Way Galaxy has a linear diameter of 1.0 AU, what will its angular diameter be as seen from Earth? Assume the distance to the center of the galaxy is 8.2 kpc. (*Hint:* Use the small-angle formula, Chapter 3.)

Learning to Look

1. Why does the galaxy shown below have so much dust in its disk? How big do you suppose the halo of that galaxy really is?

NASA/Hubble Heritage Team/STScI/AURA

2. Why are the spiral arms in the galaxy below blue? What color would the halo be if it were bright enough to see in this photo?

NASA/Hubble Heritage Team and A. Riess, STScI

Great Debates

1. *Which Formed First?* The Milky Way Galaxy has a bulge, a disk with arms, and a visible halo. Which part formed first, second, and third? How do you know? Repeat the process but add in the fourth part—the dark halo.
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

2. *Are You Made of Dark Matter?* If dark matter is in the Milky Way Galaxy, the solar system is in the Milky Way Galaxy, and you are part of the solar system, does part of you contain dark matter? How about the space around you on Earth and in the solar system? How do you know?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

3. *Solar Systems and Galaxies.* A child tells you that we live in a galaxy called the solar system. You know our solar system lives with many other star systems in a galaxy called the Milky Way Galaxy. Do you correct the child on the spot, do you find the parent and ask to correct the child, do you ignore the child's mistake, or what do you do? What if the person were older, say a co-worker or your boss, and the gaffe occurred in an e-mail to a group of people including a client? Then what do you do?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

4. *Population III Stars.* Population III stars are a hypothetical family of stars that are metal free, very massive, and hot. These stars are thought to have formed when the universe was young. These stars have not yet been observed directly but are inferred to exist. Should students learn about Population III stars? Given that not everything can be included in a textbook, what information would you remove to make room for this new subject?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

5. *MWG Ambassador.* To best represent our home galaxy, which celestial object in the Milky Way Galaxy should be our ambassador? Explain why you choose this object.
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Self-Sustaining Star Formation

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 16: Astronomical Distance Scales (Cepheid Variables)

Determining distances to astronomical objects is both crucial and very difficult. Without knowing the distance to an object, you have no way of translating its angular size into its true size or its apparent brightness into its true energy output (luminosity). You don't really know the nature of an object unless you know its distance.

In previous chapters you learned that scientists extended humanity's reach into the cosmos first by determining distances inside the solar system using Earth's size as a parallax baseline. Having measured the solar system, scientists could survey distances to nearby stars using the radius of Earth's orbit, 1 AU, as the baseline of much larger parallax triangles. But stellar parallax measurements are not practical for distances greater than a few hundred parsecs, even from current space observatories, because of the tiny parallax shifts involved. To reach greater distances, astronomers use the inverse square law. You may recall that this method depends on knowing that the distant object has the same luminosity as the nearer object, so their brightness ratio equals the inverse square of their distance ratio.

Astronomers realized that stars with the exact same spectral types should be identical in luminosity because they have the same surface temperature and diameter. Thus, any relatively nearby star with a distance known directly by parallax can be used as a "standard candle" for inverse-square calculations of distances of more distant stars of the same type. (Note that the term *standard candle* must date back to before common use of electric lightbulbs!) As you also learned in an earlier chapter, astronomers found that correction had to be made to the basic inverse-square method for the effects of interstellar dust extinction as they explored the farthest reaches of the Milky Way Galaxy.

In this chapter you learned that a particular type of pulsating variable star named Cepheids are an especially useful type of standard candle because two Cepheid stars with the same pulsation period are identical in luminosity. You don't have to laboriously take a detailed spectrum of a faint, distant Cepheid to determine its type; you only need to monitor its brightness over the span of a few days or weeks and determine its period. Also, Cepheids are giant stars, so they can be detected at great distances. Of course, as you learned in this chapter, some nearby Cepheids had to have their distances determined directly, not by the inverse-square method, before the period–luminosity relationship could be calibrated.

Caution: Reality is a little more complicated than this simple scheme. The first astronomers to measure distances using Cepheid variables as standard candles did not realize there are two different kinds of Cepheids; ones near the plane of the Milky Way have different chemical compositions than ones in the halo of the galaxy. Thus, the two types have different internal structure and somewhat different period–luminosity relationships. Once that wrinkle was worked out, astronomers were able to use Cepheids to determine distances throughout the Milky Way and even to nearby galaxies.

Section 3 of Virtual Astronomy Lab 16, "Astronomical Distance Scales," lets you practice determining periods, luminosities, and distances of Cepheid variable stars. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 18: Active Galactic Nuclei

Is there really a supermassive black hole at the center of our galaxy? As you learned in this chapter, the evidence is now unequivocal: Motions of gas clouds and stars in the galactic nucleus prove that more than 4 million solar masses lie within a region not much larger than our solar system centered on the unique radio source named Sagittarius A* (Sgr A*). A black hole is the only hypothesis that matches all the observations. You will discover in the next chapter that, as tremendous as the Milky Way's central black hole may seem, it is a puppy compared with the black holes in the nuclei of other galaxies.

Section 3 of Virtual Astronomy Lab 18, "Active Galactic Nuclei," lets you repeat the analysis that leads to the conclusion about the nature of Sgr A*. An animation called an orbit builder, along with the final exercise in the lab, provides simulated observations of two stars that orbit close to the center of the galaxy. Using those data you can calculate the mass of the object the stars are orbiting. You will discover that, although the object has a small diameter, it has a huge mass and must be a supermassive black hole. Also, note that you get two estimates of the mass of the galaxy's central object that do not agree exactly. They are based on "observations," and realistically contain some natural uncertainty.

Whenever you make a measurement, you must expect uncertainty. Like a scientist in the same situation, you need to combine separate estimates of a quantity to get a higher-quality final result. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 17: Evidence for Dark Matter

If dark matter is invisible, why do astronomers talk about it all the time? In fact, it is invisible to observations made using light and other types of electromagnetic radiation. That must mean that dark matter is composed of particles that don't interact with photons or with charged particles like electrons and protons. Nevertheless, astronomers know that dark matter exists because it exerts gravitational influence on normal matter, and the effects of that can be detected easily. For example, as you learned in this chapter, our Milky Way Galaxy rotates too fast given the amount of visible matter it contains, and the discrepancy is large. From that discrepancy astronomers can calculate how much dark matter must be present in the Milky Way, and there is much more of it than there is visible matter. Astronomers "see" dark matter by the influence of its gravity. You can be sure it exists.

Astronomers can observe stars and clouds of gas made of normal matter all though the Milky Way Galaxy and, using the Doppler effect, measure the orbital velocity of that matter as it circles the center of the galaxy A plot of orbital velocity versus distance from the center of the galaxy is know as a "rotation curve." (Never mind that it would be more correct to say that the stars and clouds are revolving around the center, not rotating.) If most of the matter in the galaxy were located near the center, you would expect the rotation curve to drop at large distances from the center because the further a star is from that central mass, the less strong would be the gravity pulling it inward and the lower its orbital velocity. That would make sense; it is the pattern seen on a much smaller scale in the solar system. Instead, observations show stars and gas clouds in the outer parts of the galaxy do not orbit slower than the material closer in—the actual rotation curve is described as "flat." That must mean there is some extra invisible mass spread through the galaxy exerting gravitational influence.

The evidence for dark matter is very strong. It really is there even if it is invisible, and that must mean it is made up of some undiscovered kind of particle that is all around us but the existence of which was not realized until astronomers noticed its gravitational effects.

Section 1 of Virtual Astronomy Lab 17, "Evidence for Dark Matter," guides you through making calculations of the amount of dark matter present in the Milky Way based on its rotation curve. Remember to be careful about units. To make any calculation properly, you have to think about the units you are using, whether they are miles per gallon or light years per parsec. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

This artist's impression shows the core of a distant galaxy. A supermassive black hole at the center is absorbing matter and ejecting a powerful jet to the upper right. Most large galaxies have a massive black hole at their center, but only a few are active at any one time.

ESO/M. Kornmesser

18

Galaxies: Normal and Active

Guidepost

Our Milky Way Galaxy is only one of the many billions of galaxies visible in the sky. This chapter will expand your horizon to discuss the different kinds of galaxies, their complex histories, and violent eruptions. Here you can expect answers to four important questions:

- **How do astronomers know what galaxies are like?**
- **Do other galaxies contain dark matter and supermassive black holes, as does our own galaxy?**
- **Why are there different kinds of galaxies?**
- **Why do some galaxies produce tremendous eruptions?**

As you begin studying galaxies, you will discover they are classified into different types, and that will lead you to insights into how galaxies form, interact, and evolve.

A hypothesis or theory is clear, decisive, and positive, but it is believed by no one but the man who created it. Experimental findings, on the other hand, are messy, inexact things which are believed by everyone except the man who did that work.

HARLOW SHAPLEY,
THROUGH RUGGED WAYS TO THE STARS

Science fiction heroes flit effortlessly between the stars, but almost none travels between the galaxies. As you leave your home galaxy, the Milky Way, behind, you will voyage out into the depths of the universe, out among the galaxies, into space so deep it is unexplored even in fiction.

18-1 The Family of Galaxies

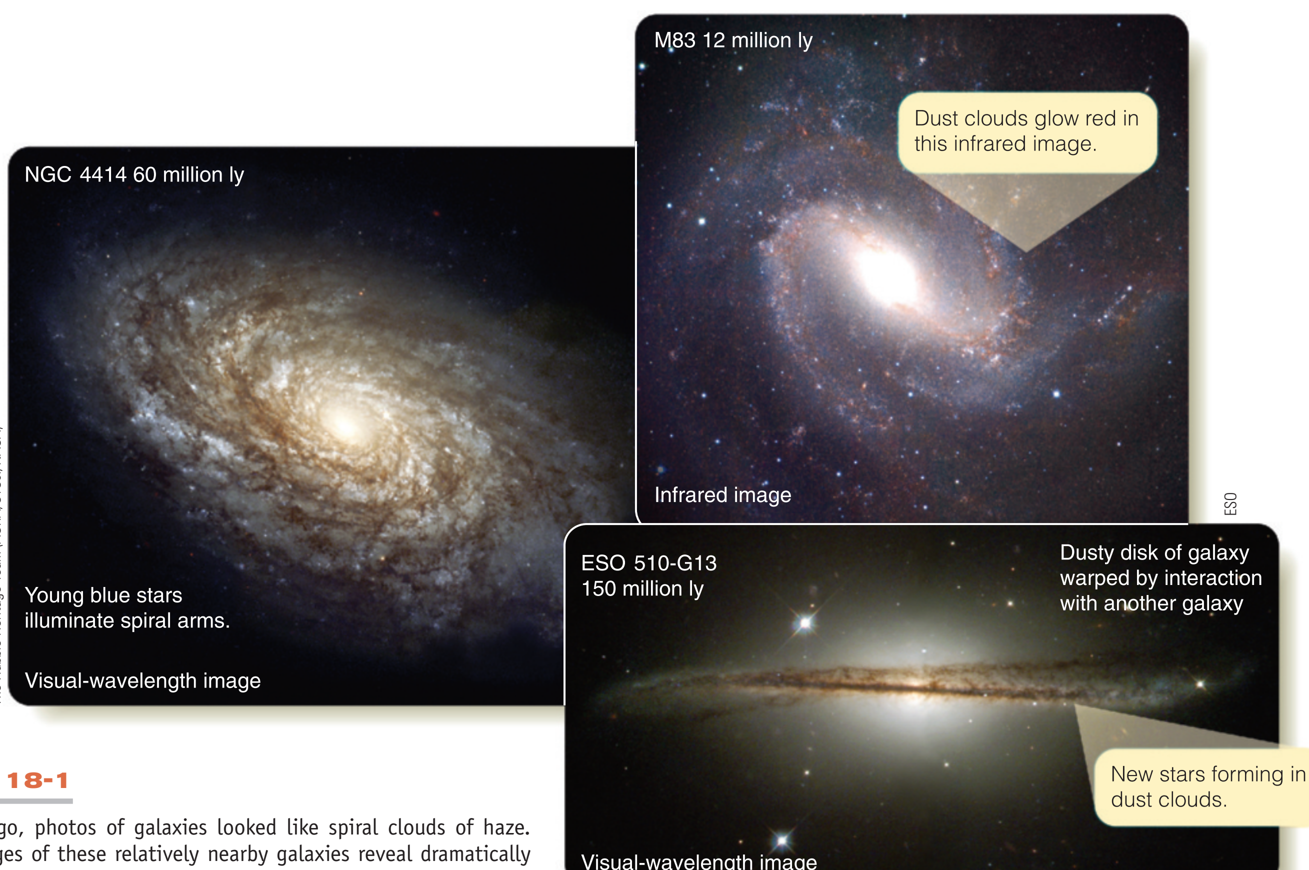
Astronomy books often include pictures of spiral galaxies. Like movie stars, spiral galaxies get a lot of attention because they are beautiful (■ Figure 18-1); but many galaxies are nearly featureless clouds of stars, and others are distorted muddles of gas and dust. You must begin by sorting out this jumble of galaxies.

The Shapes of Galaxies

Astronomers classify galaxies according to their shapes in photographs made at visual wavelengths using a system developed by Edwin Hubble in the 1920s. Such systems of classification are a fundamental technique in science (**How Do We Know? 18-1**).

Read **Galaxy Classification** on pages 398–399 and notice three important points and four new terms that describe the main types of galaxies:

1. Many galaxies have no disk, no spiral arms, and almost no gas and dust. These *elliptical galaxies* range from huge giants to small dwarfs.
2. Disk-shaped galaxies usually have spiral arms and contain gas and dust. Many of these *spiral galaxies* have a central region shaped like an elongated bar and are called *barred spiral galaxies.* A few disk galaxies contain little gas and dust.
3. *Irregular galaxies* are generally shapeless and tend to be rich in gas and dust.



The Hubble Heritage Team (AURA/STScI/NASA)

ESO

Hubble Heritage Team, AURA/STScI/NASA

■ **Figure 18-1**

A century ago, photos of galaxies looked like spiral clouds of haze. Modern images of these relatively nearby galaxies reveal dramatically beautiful objects filled with newborn stars and clouds of gas and dust.

Classification in Science

What does classification tell a scientist? Classification is one of the most basic and most powerful of scientific tools. Establishing a system of classification is often the first step in studying a new aspect of nature, and it can produce unexpected insights.

Charles Darwin sailed around the world from 1831 to 1836 with a scientific expedition aboard the ship HMS *Beagle*. Everywhere he went, he studied the living things he saw and tried to classify them. For example, he classified different types of finches he saw on the Galapagos Islands based on the shapes of their beaks. He found that those that fed on seeds with hard shells had thick, powerful beaks, whereas those that picked insects out of deep crevices had long, thin beaks. His classifications of these and other animals led him to think about how natural selection shapes creatures to survive in their environment, which led him to understand how living things evolve.

Years after Darwin's work, paleontologists classified dinosaurs into two orders, lizard-hipped and bird-hipped dinosaurs. This classification, based on the shapes of dinosaur hip joints, helped the scientists understand patterns of evolution of dinosaurs. It also led to the conclusion that modern birds, including the finches that Darwin saw on the Galapagos, evolved from dinosaurs.

Astronomers use classifications of galaxies, stars, moons, and many other objects to help them see patterns, trace relationships, and generally make sense of the astronomical world. Whenever you encounter a scientific discussion, look for the classifications on which it is based. Classifications are the orderly framework on which much of science is built.


M. Seeds

The careful classification of living things has revealed that the birds, including this flamingo, are descended from dinosaurs.

It is surprisingly difficult to figure out what proportions of galaxies are elliptical, spiral, or irregular. In catalogs of galaxies, about 70 percent are spiral, but that is because spiral galaxies are luminous and easy to notice in that they contain hot, bright stars and clouds of ionized gas (**How Do We Know? 18-2**). Most ellipticals are fainter and harder to notice. Small galaxies such as dwarf ellipticals and dwarf irregulars are actually very common, but they are hard to detect. From careful statistical studies, astronomers estimate that ellipticals are actually more common than spirals, and that irregulars make up only about 25 percent of all galaxies.

How many galaxies are there? Long exposure images of small areas on the sky are called deep fields because they detect the most distant galaxies deep in space (■ Figure 18-2). Such images show that galaxies carpet the sky like leaves on the forest floor, and as larger telescopes are built, astronomers will see even more galaxies. At least 300 billion are detectable with today's telescopes.

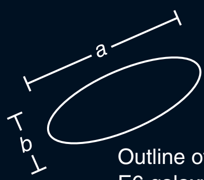
© Cengage Learning 2014; image: R. Williams and the Hubble Deep Field Team, STScI, NASA

■ Figure 18-2

An apparently empty spot on the sky only 1/30 the diameter of the full moon contains over 1500 galaxies in this extremely long time exposure known as the Northern Hubble Deep Field. Only four stars are visible in this image; they are sharp points of light with diffraction spikes produced by the telescope optics. Presumably the entire sky is similarly filled with galaxies.

1 **Elliptical galaxies** are round or elliptical, contain no visible gas and dust, and lack hot, bright stars. They are classified with a numerical index ranging from 1 to 7; E0s are round, and E7s are highly elliptical. The index is calculated from the largest and smallest diameter of the galaxy used in the following formula and rounded to the nearest integer.

$$\frac{10(a - b)}{a}$$

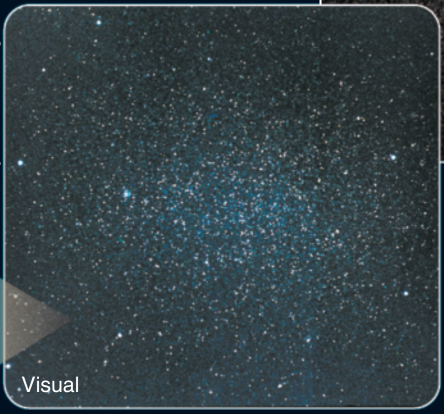


Outline of an E6 galaxy

Visual-wavelength image

NOAO/AURA/NSF

M87 is a giant elliptical galaxy classified E1. It is a number of times larger in diameter than our own galaxy and is surrounded by a swarm of over 500 globular clusters.

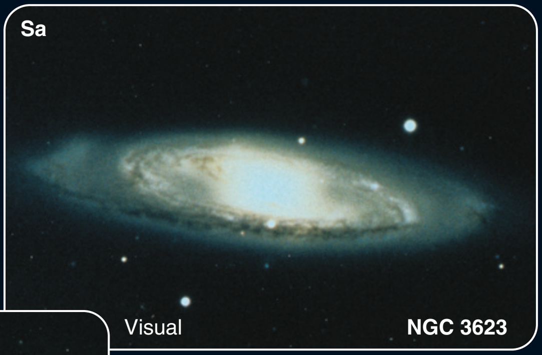


Visual

Australian Astronomical Observatory/David Malin Images

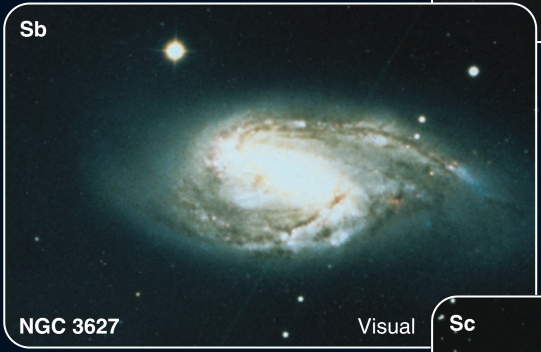
The Leo 1 dwarf elliptical galaxy is not many times bigger than a globular cluster.

2 **Spiral galaxies** contain a disk and spiral arms. Their halo stars are not visible, but presumably all spiral galaxies have halos. Spirals contain gas and dust and hot, bright O and B stars, as shown at right and below. The presence of short-lived O and B stars alerts us that star formation is occurring in these galaxies. Sa galaxies have larger nuclei, less gas and dust, and fewer hot, bright stars. Sc galaxies have small nuclei, lots of gas and dust, and many hot, bright stars. Sb galaxies are intermediate.



Sa — Visual — NGC 3623

Australian Astronomical Observatory/David Malin Images

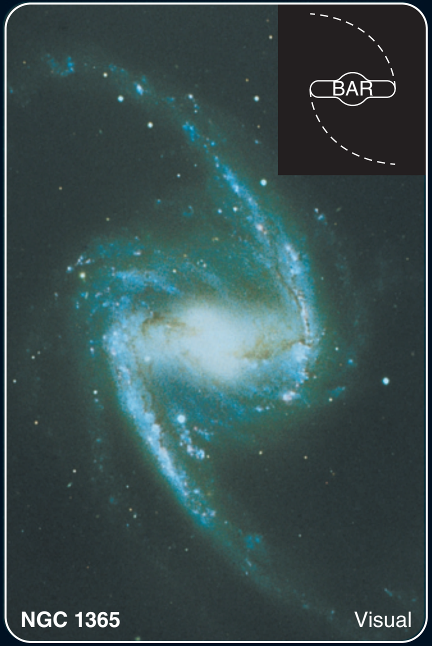


Sb — NGC 3627 — Visual

Australian Astronomical Observatory/David Malin Images

Sc — NGC 2997 — Visual

Australian Astronomical Observatory/David Malin Images



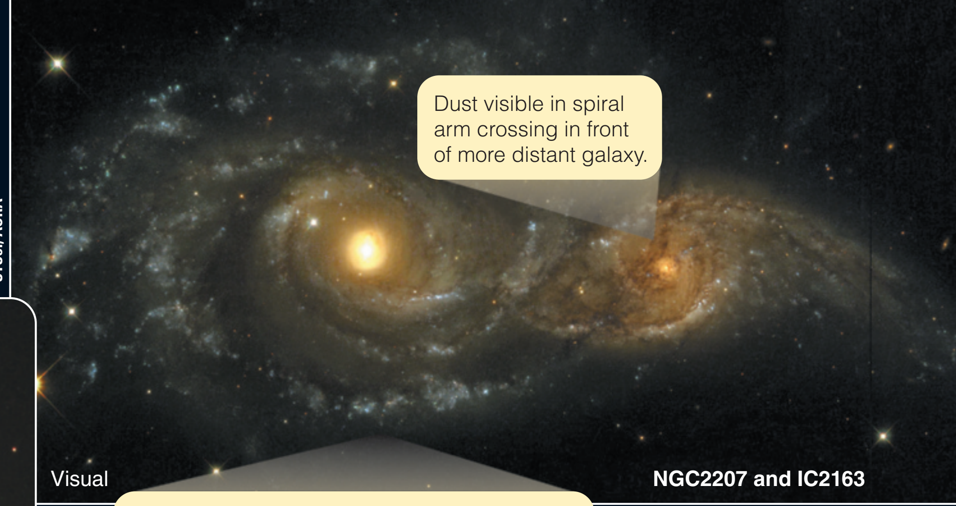
NGC 1365 — Visual

Australian Astronomical Observatory/David Malin Images

2a Roughly 2/3 of all spiral galaxies are **barred spiral galaxies** classified SBa, SBb, and SBc. They have an elongated nucleus with spiral arms springing from the ends of the bar, as shown at left. Our own galaxy is a barred spiral.

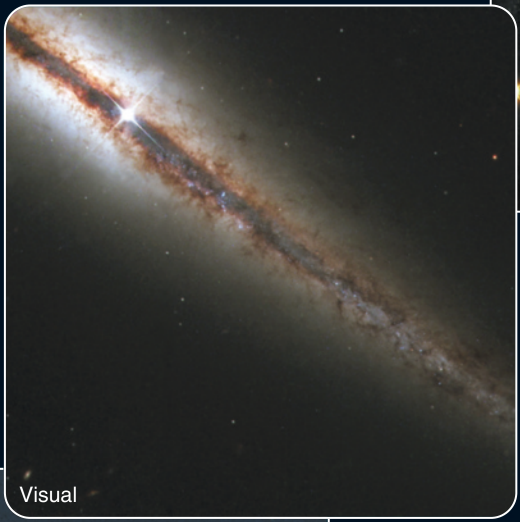
2b Some disk galaxies are rich in dust, which is concentrated along their spiral arms. NGC4013, shown below, is a galaxy much like ours, but seen edge-on its dust is dramatically apparent.

NASA, The Hubble Heritage Team, STScI, AURA

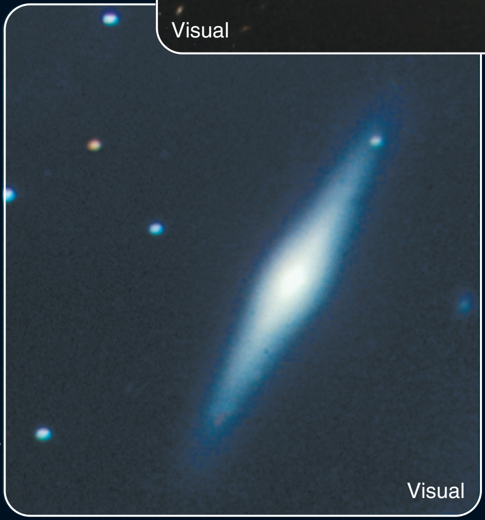


Dust in spiral galaxies is most common in the spiral arms. Here the spiral arms of one galaxy are silhouetted in front of a more distant galaxy.

NASA and The Hubble Heritage Team (STScI/AURA)

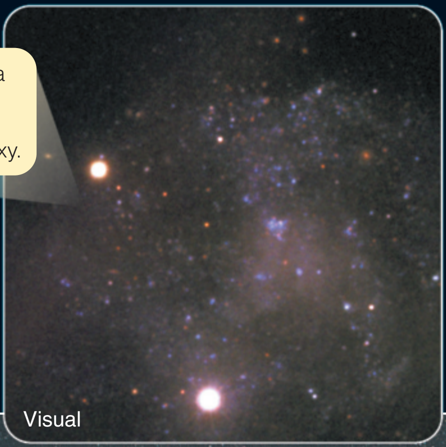


Rudolph E. Schild



2c Galaxies with an obvious disk and nuclear bulge but no visible gas and dust and no hot, bright stars are classified as S0 (pronounced “Ess Zero”). Compare this galaxy with the edge-on spiral above.

The galaxy IC4182 is a dwarf irregular galaxy only about 4 million parsecs from our galaxy.



G.J.Jacoby, M.J.Pierce/AURA/NOAO/NSF

3 **Irregular galaxies** (classified Irr) are a chaotic mix of gas, dust, and stars with no obvious nuclear bulge or spiral arms. The Large and Small Magellanic Clouds are visible to the unaided eye as hazy patches in the southern hemisphere sky. Telescopic images show that they are irregular galaxies that are interacting gravitationally with our own much larger galaxy. Star formation is dramatic in the Magellanic Clouds. The bright pink regions are emission nebulae excited by newborn O and B stars. The brightest nebula in the Large Magellanic Cloud is called the Tarantula Nebula.

© R. J. Dufour, Rice University

NOAO/AURA/NSF

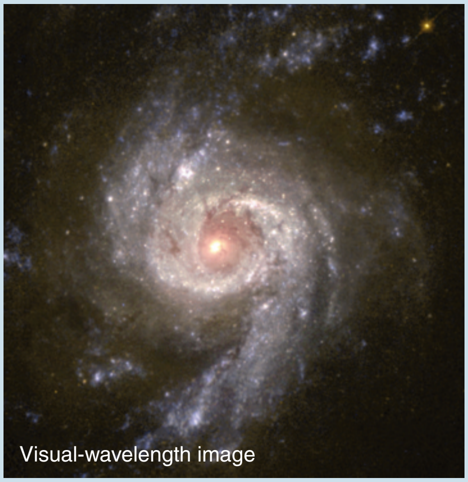
How Do We Know? | 18-2

Selection Effects

How can selecting what to study be misleading? Scientists must design their research projects with great care. Biologists studying insects in the rain forest, for example, must choose which ones to catch. Because they can't catch every insect they see, they might decide to catch and study any insect that is red. But if they are not careful as they design their research plan, a selection effect could bias their data and lead them to incorrect conclusions without their ever knowing it.

To see how this could happen, suppose you needed to measure the speed of cars on a highway. There are too many cars to measure every one, so you might decide to measure only red cars. It is quite possible that this selection criterion will mislead you because people who buy red cars may be more likely to be younger and drive faster. Should you measure only brown cars? No, because older, more sedate people might tend to buy brown cars. Only by very carefully designing your experiment can you be certain that the cars you measure are traveling at representative speeds.

Astronomers understand that what you see through a telescope depends on what you notice, and that is powerfully influenced by selection effects. The biologists in the rain forest, for example, should not catch and study only red insects. Often, the most brightly colored insects are poisonous or at least taste bad to predators. Catching only red insects could lead the scientists to false conclusions about the kinds of insects that live in the forest.



Hubble Heritage Team/STScI/AURA/NASA

Things that are bright and beautiful, such as spiral galaxies, may attract a disproportionate amount of attention. Scientists must be aware of such selection effects.

SCIENTIFIC ARGUMENT

What color are galaxies?

Scientific arguments must be based on evidence, and evidence means observations. But you need to be careful and analyze even the simplest observations with care. Different kinds of galaxies have different colors, depending mostly on how much gas and dust they contain. If a galaxy contains large amounts of gas and dust, it probably contains lots of young stars, and a few of those young stars will be massive, hot, luminous O and B stars. They will produce most of the light and give the galaxy a distinct blue tint. In contrast, a galaxy that contains little gas and dust will contain few young stars. It will lack O and B stars, and instead the most luminous stars in such a galaxy will be red giants and supergiants. They will give the galaxy a red tint. Because the light from a galaxy is a blend of the light from billions of stars, the colors are only tints. Nevertheless, the most luminous stars in a galaxy determine the overall color. From this you can conclude that elliptical galaxies tend to be red, and the disks of spiral galaxies tend to be blue.

Now create your own scientific argument and analyze a different kind of observation. **Why are most cataloged galaxies spiral in spite of the fact that the most common kind of galaxy is elliptical?**

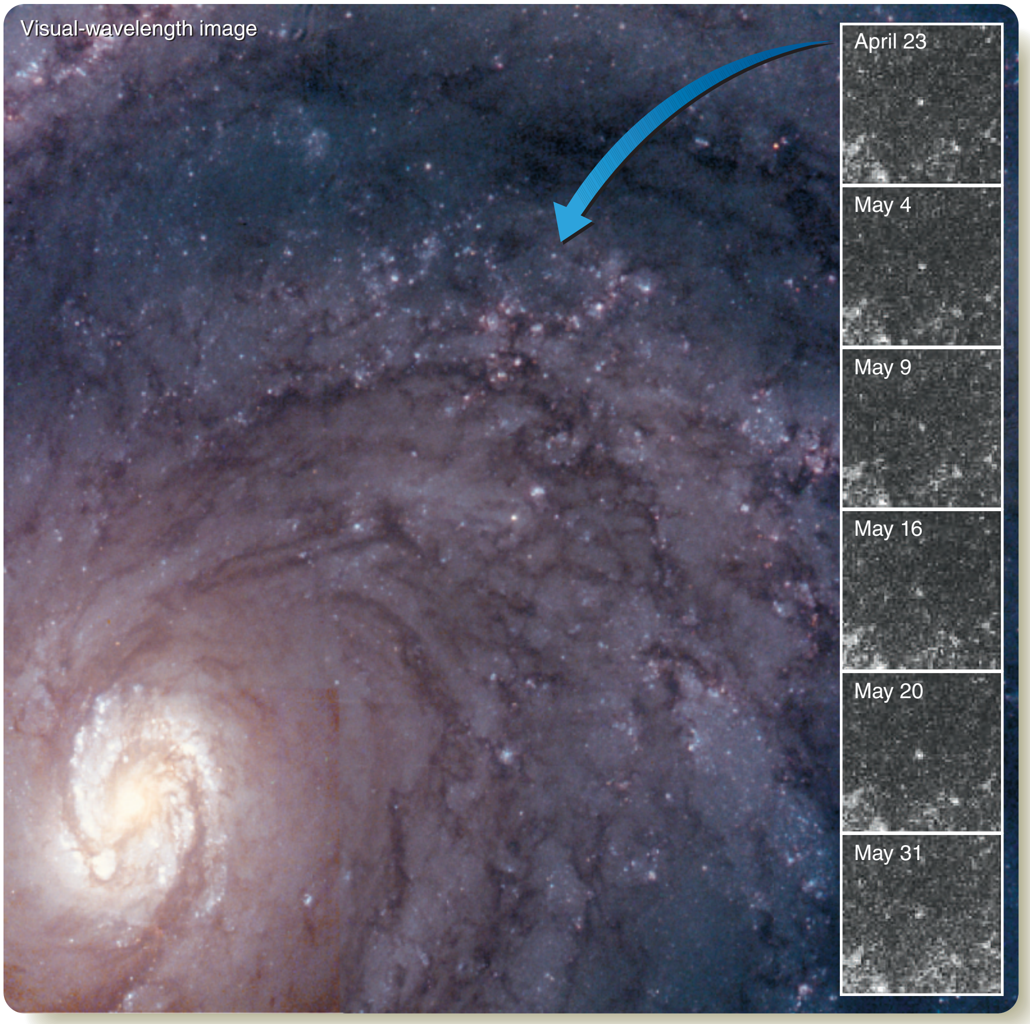
18-2 Measuring the Properties of Galaxies

WHAT ARE THE PROPERTIES OF GALAXIES? What are their diameters, luminosities, and masses? Just as in your study of stellar characteristics (see Chapter 13), the first step in your study of galaxies is to find out how far away they are. Once you know a galaxy's distance, its size and luminosity are relatively easy to find. Later in this section, you will see that, just as for stars, finding the masses of galaxies is more difficult.

Distance

The distances to galaxies are so large that it is not convenient to measure them in light-years, parsecs, or even kiloparsecs. Instead, astronomers use the unit **megaparsec (Mpc),** or 1 million pc. One Mpc equals 3.26 million ly, or approximately 3×10^{19} km (2×10^{19} miles).

To find the distance to a galaxy, astronomers must search among its stars, nebulae, and star clusters for familiar objects whose luminosity they know. Such objects are called **distance indicators** because they can be used to find the distance to a galaxy. Astronomers often refer to them as **standard candles.** If you can find a standard candle in a galaxy, you can judge its distance.



© Cengage Learning 2014; image: J. Trauger, JPL and NASA

■ **Figure 18-3**

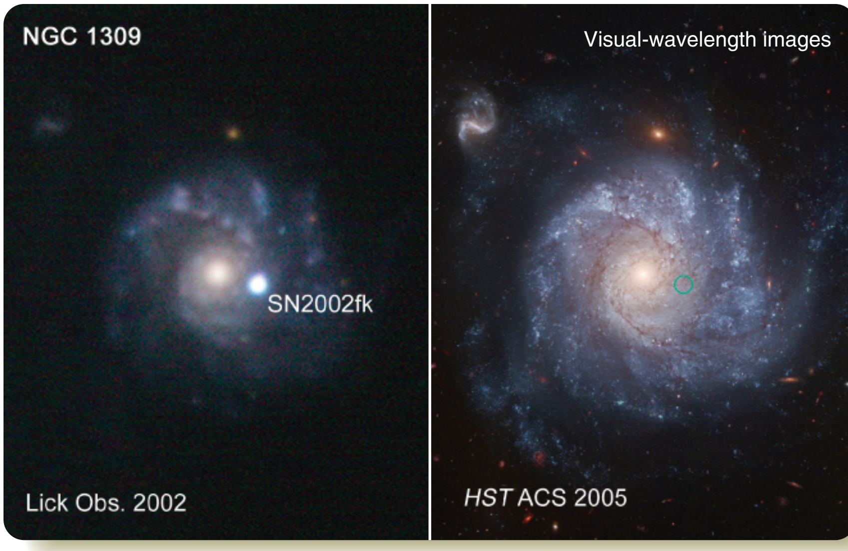
The vast majority of spiral galaxies are too distant for Earth-based telescopes to detect Cepheid variable stars. The Hubble Space Telescope, however, can locate Cepheids in some of these galaxies, as it has in the bright spiral galaxy M100. From a series of images taken on different dates, astronomers can locate Cepheids (inset), determine the period of pulsation, and measure the average apparent brightness. They can then deduce the distance to the galaxy—16 Mpc (52 million ly) for M100.

Cepheid variable stars are reliable distance indicators because their period is related to their luminosity. The period–luminosity relation has been calibrated (see Figure 17-4 and also How Do We Know? 17-1), so you can use the period of the star's variation to find its absolute magnitude. Then, by comparing its absolute and apparent magnitudes, you can find its distance. ■ Figure 18-3 shows a galaxy in which the Hubble Space Telescope detected Cepheids.

Even with the Hubble Space Telescope, Cepheids are not visible in galaxies much beyond 30 Mpc (100 million ly), so astronomers must search for less common but brighter distance indicators and calibrate them using nearby galaxies that contain both indicators. For example, by studying nearby galaxies with distances known from Cepheids, astronomers have found that the brightest globular clusters have absolute magnitudes of about -10. If you found globular clusters in a more distant galaxy, you could assume that the brightest of the globular clusters have absolute magnitudes of -10 and use that information to calculate the distance.

Astronomers can calibrate type Ia supernovae, those produced by the collapse of a white dwarf, because the white dwarf always collapses at the same mass limit, so the explosions reach approximately the same maximum luminosity. When type Ia supernovae occur in galaxies with distances that are known from Cepheid variables and other distance indicators, astronomers can find the absolute magnitude of the supernovae at peak brightness. An example is shown in ■ Figure 18-4. When astronomers see a type Ia supernova in a more distant galaxy, they observe the apparent magnitude of the supernova at maximum and compare that with the known absolute magnitude these supernovae reach at maximum to find the distance to the galaxy. As you will see in the next chapter, this is a critical calibration in modern measurements of the size and age of the universe.

Notice how astronomers use calibration to build a **distance scale** reaching from the nearest galaxies to the most distant visible galaxies. Often astronomers refer to this as the "distance pyramid" or the "distance ladder" because each step depends on the steps below it. The most dependable step is the Cepheid variable stars, but notice that Cepheid distance indicators depend on astronomers' understanding of the luminosities of the stars in the H–R diagram, and that rests on an even lower step in the ladder—measurements of the parallax of stars. Stellar parallaxes in turn depend on measuring the size of Earth's orbit around the sun, which finally rests on the bottom step, measuring the size of Earth itself. The distance ladder ultimately connects the size of Earth to the most distant galaxies in the universe.



Lick Observatory: W. Li and A.V. Filippenko

NASA, ESA, The Hubble Heritage Team, STScI/AURA and A. Riess, STScI

Figure 18-4

While dinosaurs roamed Earth, a white dwarf in the galaxy NGC 1309 collapsed and exploded as a type Ia supernova, and the light from that explosion reached Earth in 2002. Astronomers found Cepheid variable stars in the galaxy, so they can determine that it is 100 million light-years away, and that allowed them to find the absolute magnitude of the supernova at its brightest. By combining observations of many supernovae, astronomers were able to calibrate type Ia supernovae as distance indicators.

The most distant visible galaxies are over 3000 Mpc (10 billion ly) away, and at such distances you see an effect akin to time travel. When you look at a galaxy that is millions of light-years away, you do not see it as it is now but as it was millions of years ago when its light began the journey toward Earth. When you look at a distant galaxy, you look back into the past by an amount called the **look-back time,** a time in years equal to the distance in light-years the light from the galaxy traveled to reach Earth.

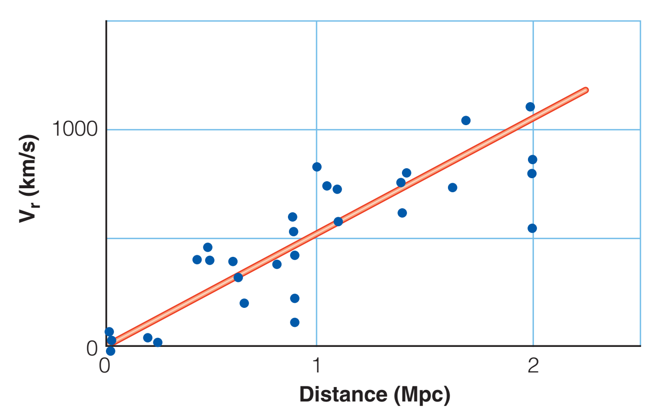
The look-back time to nearby objects is usually not significant. For example, the look-back time across a football field is a tiny fraction of a second. The look-back time to the moon is 1.3 seconds, to the sun only 8 minutes, and to the nearest star a bit over 4 years. The Andromeda Galaxy has a look-back time of over 2 million years, but that is a mere eyeblink in the lifetime of a galaxy. If you look at more distant galaxies, the look-back time becomes an appreciable part of the age of the universe. In the next chapter, you will learn about evidence that the universe began almost 14 billion years ago. When you look at the most distant visible galaxies, you are looking back over 10 billion years to a time when the universe was significantly different. The look-back time becomes an important factor as you begin to think about the origin and evolution of galaxies.

The Hubble Law

Although astronomers must work carefully to measure the distance to a galaxy, they often estimate such distances using a simple relationship that was first noticed at about the same time astronomers were beginning to understand the nature of galaxies.

Early in the 20th century astronomers were able to record the spectra of the brighter galaxies, and most had redshifts as if they were receding. The faintest galaxies had the largest redshifts. In the 1920s, astronomers Edwin Hubble and Milton Humason were able to measure the distances to a number of galaxies using Cepheid variable stars, and in 1929, they published a graph that plotted the apparent velocity of recession versus distance for their galaxies. The points in the graph fell along a straight line (Figure 18-5). This simple relation between apparent velocity of recession and distance is known as the **Hubble law,** and the slope of the line is called the **Hubble constant, *H*.**

The Hubble law has important implications. It is commonly interpreted to show that the universe is expanding. In the next chapter, you will study this expansion, but here you can use the Hubble law as a practical way to estimate the distance to a galaxy. As shown in **Reasoning with Numbers 18-1**, the distance to a galaxy can be found by dividing its apparent velocity of recession by the Hubble constant. This makes it relatively easy to estimate the distances to galaxies because a large telescope can photograph the spectrum of a galaxy and determine its apparent velocity of recession even when it is too distant to have visible distance indicators.



© Cengage Learning 2014

Figure 18-5

Edwin Hubble's first diagram of the apparent velocities of recession and distances of galaxies did not probe very deeply into space, and the horizontal axis was later recalibrated. These data did show, however, that the galaxies are receding from one another.

Reasoning with Numbers | 18-1

The Hubble Law

The apparent velocity of recession of a galaxy, V_r, in kilometers per second is equal to the Hubble constant, H, multiplied times the distance to the galaxy, d, in megaparsecs:

$$V_r = Hd$$

Astronomers use this as a way to estimate distance from a galaxy's apparent velocity of recession.

Example: If a galaxy has a radial velocity of 700 km/s, and H is 70 km/s/Mpc,* then the distance to the galaxy equals the velocity divided by the Hubble constant, which is

$$d = \frac{(700 \text{ km/s})}{(70 \text{ km/s/Mpc})} = 10 \text{ Mpc}$$

Notice how the units km/s cancel out to leave the distance in Mpc.

*H has the units of a velocity divided by a distance. These are usually written as km/s/Mpc, meaning km/s per Mpc.

Edwin Hubble's original measurements of H were too large because of calibration errors in his distance measurements. The most precise modern measurements of H yield a value of 70 km/s/Mpc with an uncertainty of just a few percent.

Diameter and Luminosity

Once you find the distance to a galaxy from distance indicators or the Hubble law, you can calculate its diameter and its luminosity. With a good telescope and the right equipment, you could easily photograph a galaxy and measure its angular diameter in arc seconds. If you knew the distance, you could use the small-angle formula (Reasoning with Numbers 3-1) to find its linear diameter. If you also measured the apparent magnitude of a galaxy, you could use the distance to find its absolute magnitude (Reasoning with Numbers 13-2) and from that its luminosity.

The results of such observations show that galaxies differ dramatically in size and luminosity. Irregular galaxies tend to be small, 1 to 25 percent the size of our galaxy, and of low luminosity. Although they are common, they are easy to overlook. Our Milky Way Galaxy is large and luminous compared with most spiral galaxies, though astronomers know of a few spiral galaxies that are even larger and more luminous. The largest is nearly four times bigger in diameter and about 10 times more luminous. Elliptical galaxies cover a wide range of diameters and luminosities. The largest, called giant ellipticals, are five times the size of our Milky Way, but many elliptical galaxies are very small dwarf ellipticals only 1 percent the diameter of our galaxy.

To put galaxies in perspective, you can use an analogy. If our galaxy were an 18-wheeler, the smallest dwarf galaxies would be the size of pocket-size toy cars, and the largest giant ellipticals would be the size of jumbo jets.

Of the three basic parameters that describe a galaxy, you have found two—diameter and luminosity. The third, as was the case for stars, is more difficult to measure.

Mass

Although the mass of a galaxy is difficult to determine, it is an important quantity. It tells you how much matter the galaxy contains, which in turn provides clues to the galaxy's origin and evolution.

Although there are a few ways to estimate the mass of a galaxy, the most direct measurement uses Kepler's third law of planetary motion. In Chapter 13, you used that law to find the masses of binary stars. If astronomers can measure the radius of a galaxy and the velocity with which it rotates, they can easily find the orbital periods of the stars as they orbit the galaxy. Then they can use Newton's version of Kepler's third law (Reasoning with Numbers 13-4) to calculate the mass of the galaxy. The most accurate of these studies involve observing the rotation of the galaxy at different distances from its center and drawing a graph called the **rotation curve** of the galaxy (■ Figure 18-6).

Using such a graph to find the mass of the galaxy is called the **rotation curve method.** It is the most accurate way to find the mass, but it works only for the nearer galaxies, whose rotation curves can be observed. More distant galaxies appear so small astronomers cannot measure the radial velocity at different points across the galaxy. Observations of our own galaxy and others show that the outer parts of the rotation curve do not decline to lower velocities as you would expect if most of the mass lies in the inner part of the galaxy. As in the case of the rotation curve of the Milky Way Galaxy (see Figure 17-11), this indicates that the galaxies contain large amounts of mass outside the parts visible in telescopes, perhaps in extended galactic coronas.

Measuring the masses of galaxies reveals two things. First, the range of masses is wide—from 1 million times smaller than the Milky Way Galaxy to 50 times larger. And second, galaxies contain dark matter spread through extended galactic coronae, just as does our Milky Way Galaxy.

One last region remains to be explored. What lies in the centers of galaxies? Our Milky Way Galaxy contains a

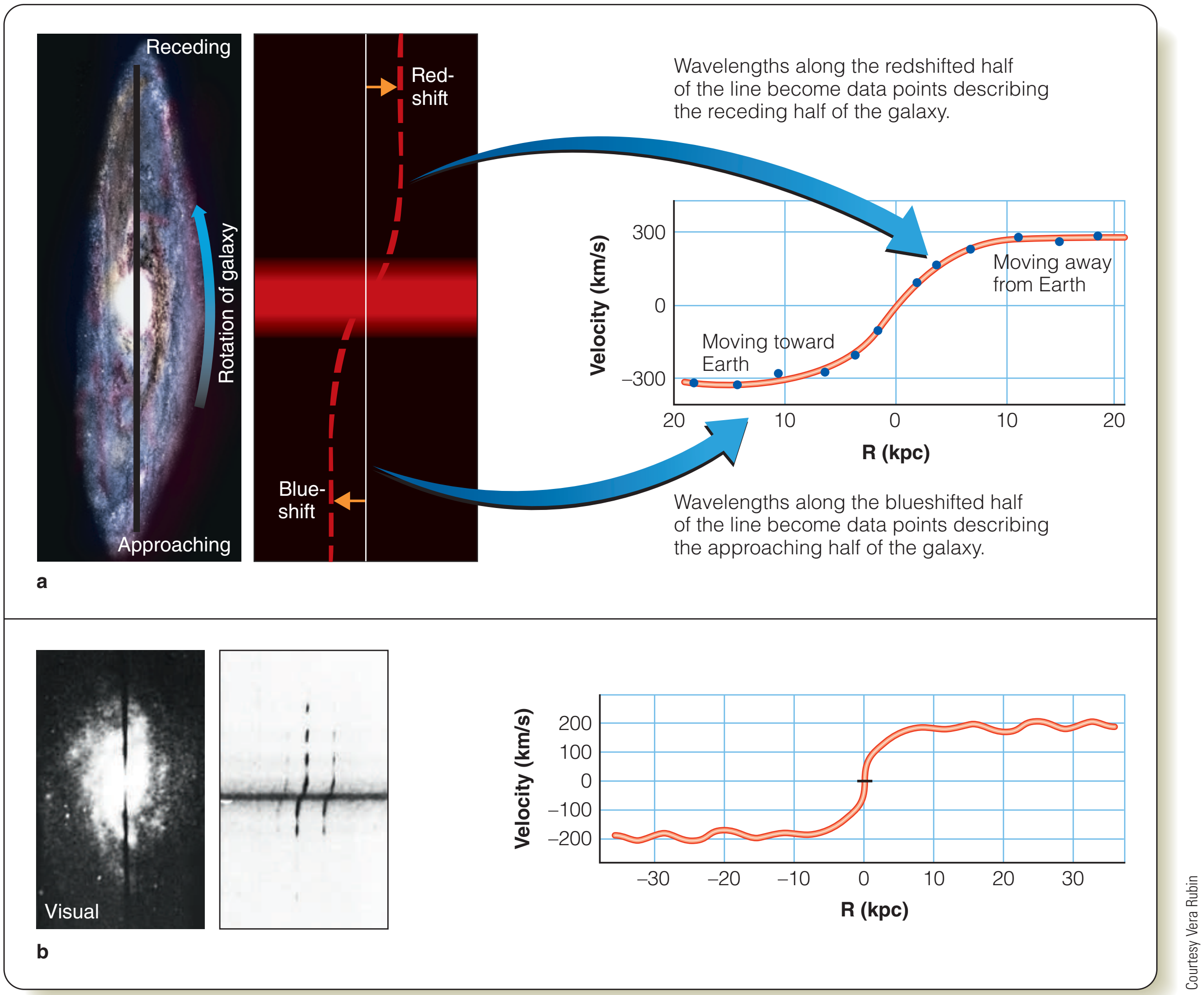


Figure 18-6

(a) In the upper panel's artwork representation, an astronomer has placed the image of the galaxy over a narrow slit so that light from the galaxy can enter the spectograph and produce a spectrum. A very short segment of the spectrum shows an emission line redshifted on the receding side of the rotating galaxy and blueshifted on the approaching side. Converting those Doppler shifts into velocities, the astronomer can plot the galaxy's rotation curve (right). (b) Real data are shown in the lower panel. Galaxy NGC 2998 is shown over the spectrograph slit, and the segment of the spectrum includes three emission lines.

4-million-solar-mass black hole at its center. Do all galaxies contain similar objects?

Supermassive Black Holes in Galaxies

Rotation curves show the motions of the outer parts of a galaxy, but it is also possible to detect the Doppler shifts of stars orbiting very close to the centers. Although these motions are not usually shown on rotation curves, they reveal something astonishing.

Measurements show that the stars near the centers of most galaxies are orbiting very rapidly. To hold stars in such small, short-period orbits, the centers of galaxies must contain millions or even billions of solar masses in a very small region. The evidence shows that the nuclei of many galaxies contain supermassive black holes. The Milky Way contains a supermassive black hole at its center (look back to Chapter 17), and evidently that is typical of galaxies.

Such a supermassive black hole cannot be the remains of a dead star. That would amount to only a few solar masses. Rather, a supermassive black hole either formed as the galaxy formed or accumulated mass over billions of years as matter sank into the center of the galaxy. Measurements show that the mass of a supermassive black hole is related to the mass of the central bulge. A galaxy with a large central bulge has a supermassive black hole whose mass is greater than the black hole in a galaxy with a small central bulge. Rare galaxies lacking central bulges usually also lack supermassive black holes. This suggests that the supermassive black holes formed long ago as the galaxies formed. Of course, matter has continued to drain into the black holes,

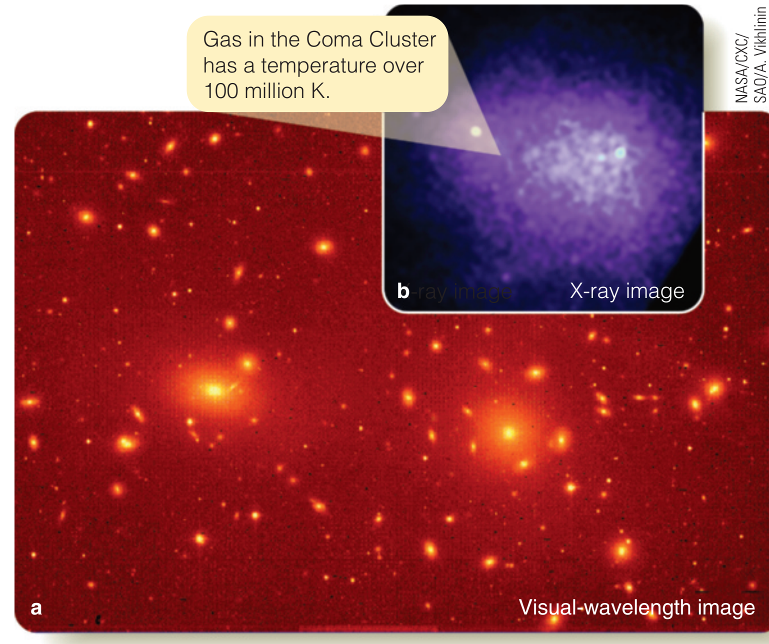
but they do not appear to have grown dramatically since they formed.

A billion-solar-mass black hole sounds like a lot of mass, but note that it is roughly 1 percent of the mass of a galaxy. The 4-million-solar-mass black hole at the center of the Milky Way Galaxy contains only one-thousandth of 1 percent of the mass of our galaxy. Later in this chapter you will discover that these supermassive black holes can produce titanic eruptions, but they still represent only a small fraction of the mass of a galaxy.

Dark Matter in Galaxies

Given the size and luminosity of a galaxy, astronomers can make a rough guess as to the amount of matter it should contain. They know how much light stars produce, and they know about how much matter there is between the stars, so it is quite possible to estimate the mass of a galaxy from its luminosity. But the measured mass of a galaxy is generally 10 times larger than that estimated from what is visible. This must mean that nearly all galaxies contain dark matter.

X-ray observations reveal more evidence of dark matter. X-ray images of galaxy clusters show that many of them are filled with very hot, low-density gas. The amount of gas present is much too small to account for the dark matter. Rather, the gas is important because it is very hot and its rapidly moving atoms have not leaked away. Evidently the gas is held in the cluster by a strong gravitational field. To have a high enough escape velocity to hold the hot gas, the cluster must contain much more matter than what astronomers can observe directly. The detectable galaxies in the Coma cluster, for instance, amount to only a small fraction of the total mass of the cluster (■ Figure 18-7).



NASA/CXC/SAO/A. Vikhlinin et al.

Gregory Bothun, University of Oregon

© Cengage Learning 2014

■ **Figure 18-7**

(a) The Coma cluster of galaxies contains at least 1000 galaxies and is especially rich in E and S0 galaxies. Two giant galaxies lie near its center. Only the central area of the cluster is shown in this image. If the cluster were visible in the sky, it would span eight times the diameter of the full moon. (b) In false colors, this X-ray image of the Coma cluster shows it filled and surrounded by hot gas. Note that the two brightest galaxies are visible in the X-ray image.

Another way to detect dark mater goes back to 1916 when Albert Einstein described gravity as a curvature of space-time. The presence of mass actually distorts space-time, and that is what you feel as gravity. He predicted that a light beam traveling through a gravitational field would be deflected by the curvature of space-time much as a golf ball is deflected as it rolls over a curved putting green. That effect has been observed and is a strong confirmation that Einstein's theories are correct.

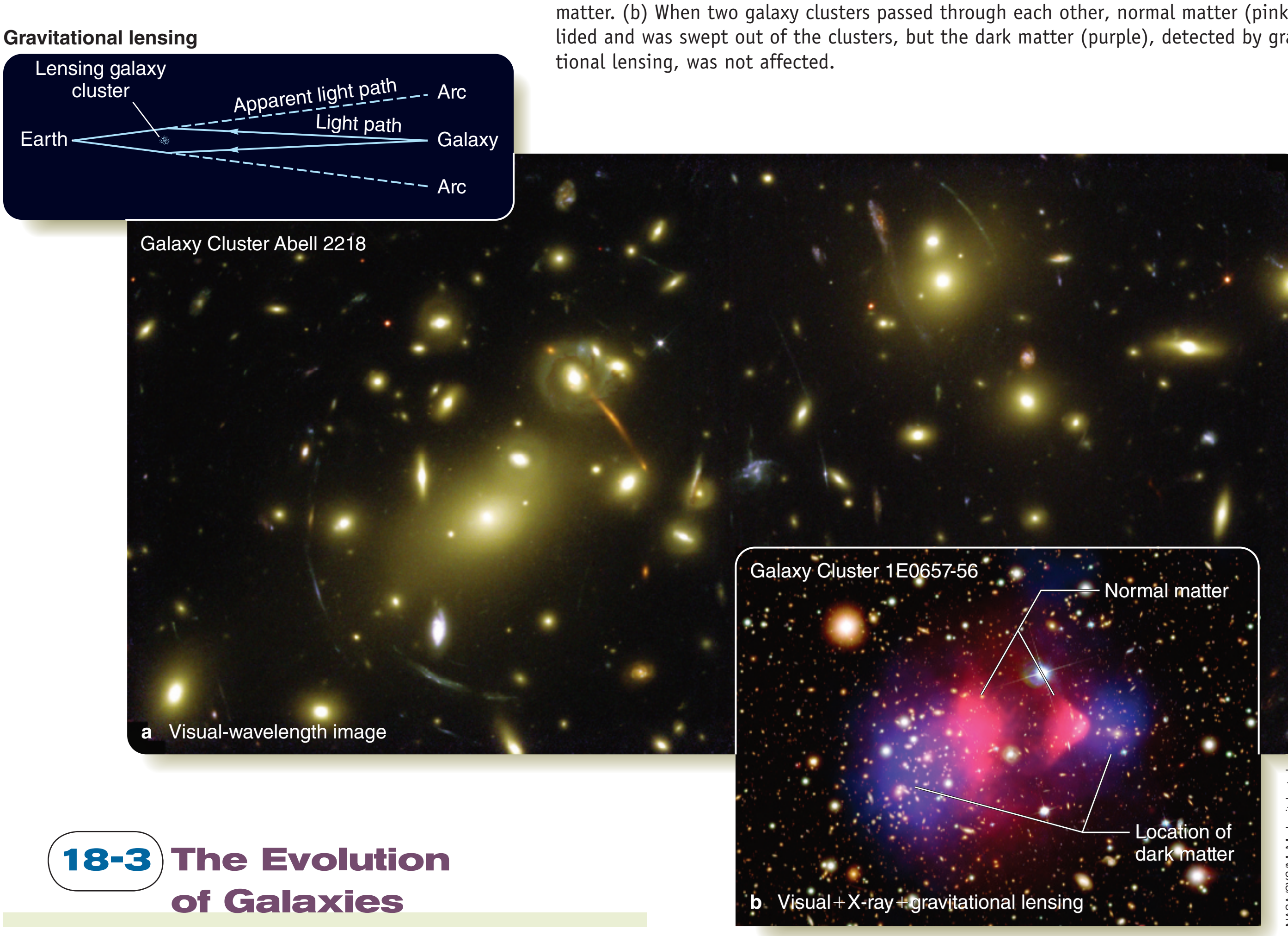
Gravitational lensing occurs when light from a distant object passes a nearby massive object and is deflected by the gravitational field. The gravitational field of the nearby object is actually a region of curved space-time that acts as a lens and deflects the passing light. Astronomers can use gravitational lensing to detect dark matter when light from very distant galaxies passes through a cluster of relatively nearby galaxies on its way to Earth and is deflected by the strong curvature. The distortion can produce multiple images of the distant galaxies and distort them into arcs. The amount of distortion depends on the mass of the cluster of galaxies (■ Figure 18-8). Observations of gravitational lensing made with very large telescopes reveal that clusters of galaxies contain far more matter than what can be seen. That is, they contain large amounts of dark matter. This confirmation of the existence of dark matter is independent of Newton's laws and gives astronomers great confidence that dark matter is real.

Theorists conclude that dark matter must be made up of some as yet undiscovered subatomic particles that do not interact with normal matter, with each other, or with light. Dark matter is detectable only through its gravitational field.

Dark matter is not an insignificant issue. Observations of galaxies and clusters of galaxies show that nearly 90 percent of the matter in the universe is dark matter. The universe you see—the kind of matter that you and the stars are made of—has been compared to the foam on an invisible ocean. Dark matter remains one of the fundamental unresolved problems of modern astronomy. You will learn more about this problem in the next chapter when you try to understand how dark matter affects the nature of the universe, its past, and its future.

■ **Figure 18-8**

(a) The gravitational lens effect is visible in galaxy cluster CL 0024+1654 as its mass bends the light of a much more distant galaxy to produce arcs that are actually distorted images of the distant galaxy. This reveals that the galaxy cluster must contain large amounts of dark matter. (b) When two galaxy clusters passed through each other, normal matter (pink) collided and was swept out of the clusters, but the dark matter (purple), detected by gravitational lensing, was not affected.



© Cengage Learning 2014

NASA, ESA, Richard Ellis (Caltech) and Jean-Paul Kneib (Observatoire Midi-Pyrenees, France)

NASA/CXC/M. Markevitch et al.

18-3 The Evolution of Galaxies

Your goal in this chapter is to build a theory to explain the evolution of galaxies. In Chapter 17 you considered the origin of our own Milky Way Galaxy; presumably, other galaxies formed similarly. But why did some galaxies become spiral, some elliptical, and some irregular? An important clue to that mystery lies in the clustering of galaxies.

Clusters of Galaxies

Single, isolated galaxies are rare. Instead, most galaxies occur in clusters containing a few to a few thousand galaxies in volumes 1 to 10 Mpc across. Our Milky Way Galaxy is a member of a small cluster, and surveys have cataloged thousands of other clusters.

For purposes of this study, you can sort clusters of galaxies into rich clusters and poor clusters. **Rich clusters** contain a thousand or more galaxies, mostly ellipticals, scattered through a volume roughly 3 Mpc (10^7 ly) in diameter. Such a cluster is nearly always condensed; that is, the galaxies are more crowded near the cluster center. At their centers, such clusters often contain one or more giant elliptical galaxies.

The Virgo cluster is an example of a rich cluster. It contains over 2500 galaxies and is about 18 Mpc (59 million ly) from our galaxy. The Virgo cluster, like most rich clusters, is centrally condensed and contains the giant elliptical galaxy M87 at its center (page 398).

Poor clusters contain fewer than a thousand (and often only a few) galaxies spread through a region that can be as large as a rich cluster. That means the galaxies are more widely separated.

Our Milky Way Galaxy is a member of a poor cluster known by the unimaginative name of the **Local Group** (■ Figure 18-9a). The total number of galaxies in the Local Group is uncertain, but it probably contains about 40 galaxies scattered irregularly through a volume roughly 1 Mpc in diameter. Of the brighter galaxies, 15 are elliptical, 4 are spiral, and 13 are irregular.

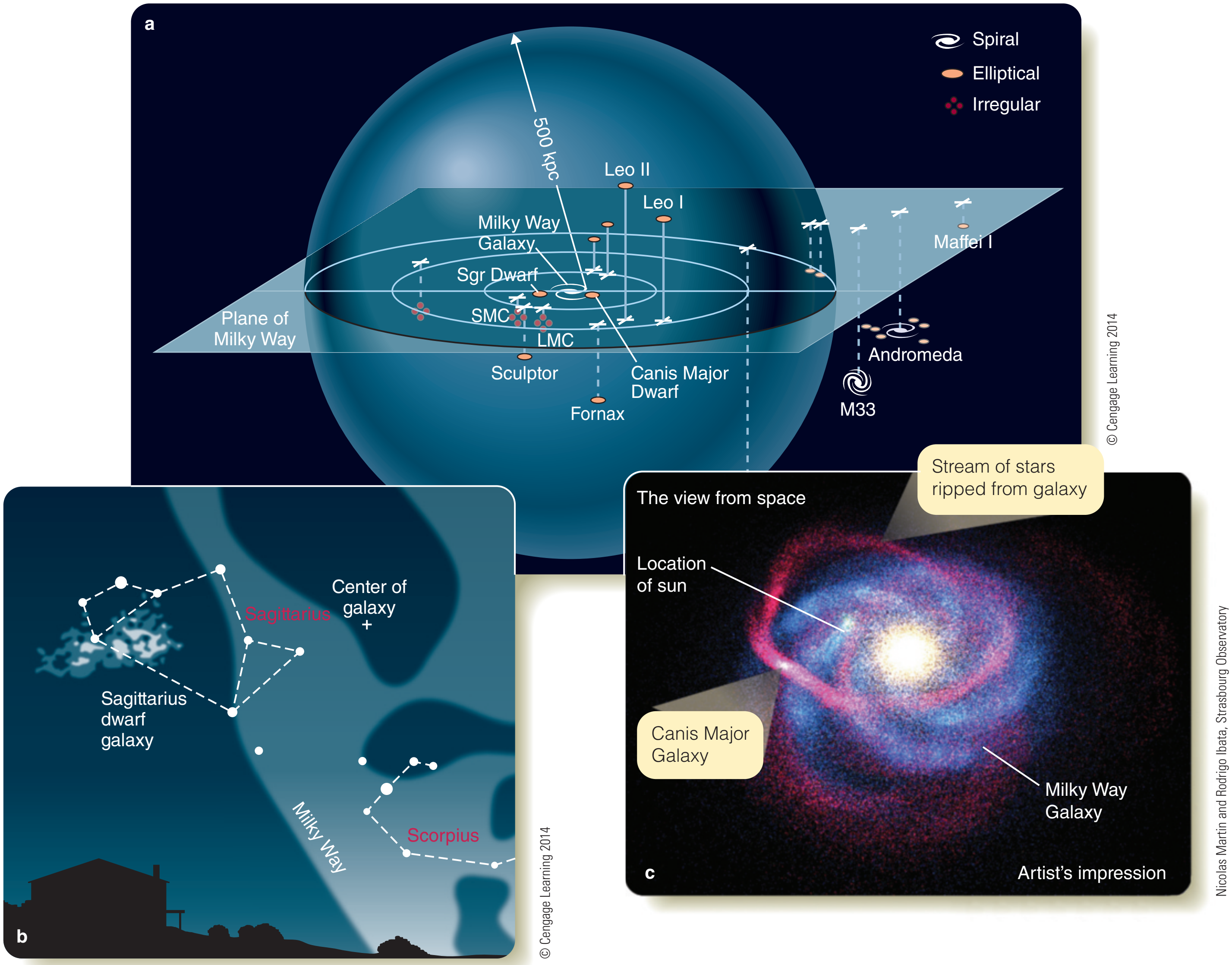


Figure 18-9

(a) The Local Group. Our galaxy is located at the center of this diagram. The vertical lines giving distances from the plane of the Milky Way are solid above the plane and dashed below. (b) The Sagittarius Dwarf Galaxy (Sgr Dwarf) lies on the other side of our galaxy. If you could see it in the sky, it would be 17 times larger than the full moon. (c) The Canis Major Dwarf Galaxy, shown in this simulation, is even closer to the Milky Way Galaxy. It is the remains of a small galaxy that is being pulled apart by our galaxy and is hidden behind the stars of the constellation Canis Major.

The total number of galaxies in the Local Group is uncertain because some galaxies lie in the plane of the Milky Way Galaxy and are difficult to detect. For example, a small galaxy known as the Sagittarius Dwarf has been found on the far side of our own galaxy, where it is almost totally hidden behind the star clouds of Sagittarius (Figure 18-9b). Even closer to the Milky Way Galaxy is the Canis Major Dwarf Galaxy (Figure 18-9c). This galaxy was found by mapping the distribution of red supergiants detected by the 2MASS infrared survey. There are certainly other small galaxies in our Local Group that have not been detected yet.

In general, rich clusters tend to contain 80 to 90 percent E and S0 galaxies and a few spirals. In other words, galaxies that are crowded together tend to be E or S0 rather than spirals. Poor clusters contain a larger percentage of spirals. Among the rare isolated galaxies that are not in clusters, over 80 percent are spirals. Somehow the environment around a galaxy helps determine its type. Astronomers suspect that collisions between galaxies are an important process.

Colliding Galaxies

Galaxies should collide fairly often. The average separation between galaxies is only about 20 times their diameter. Like two blindfolded elephants blundering about under a circus tent, galaxies should bump into each other once in a while. Stars, on the other hand, almost never collide. In the region of the galaxy near the sun, the average separation between stars is about 10^7 times

their diameters. Consequently, a collision between two stars inside a galaxy is about as likely as collision between two gnats flitting about in a football stadium.

Read **Interacting Galaxies** on pages 410–411 and notice four important points and three new terms:

1. Interacting galaxies can distort each other with tides producing *tidal tails* and shells of stars. They may even trigger the formation of spiral arms. In fact, large galaxies can even absorb smaller galaxies, a process called *galactic cannibalism.*
2. Interactions between galaxies can trigger rapid star formation.
3. Evidence left inside galaxies in the form of motions and multiple nuclei reveals that they have suffered past interactions and mergers.
4. Finally, the beautiful *ring galaxies* are understood to be bull's-eyes left behind by high-speed collisions.

Evidence of galaxy mergers is all around you. Our Milky Way Galaxy is a cannibal galaxy snacking on the nearby Magellanic Clouds. Furthermore, our galaxy's tides are pulling the Sagittarius Dwarf Galaxy apart, and the Canis Major Dwarf Galaxy has been almost completely digested as tides have pulled stars away to form great streamers wrapped around the Milky Way Galaxy (Figure 18-9c). Our galaxy has almost certainly dined on other small galaxies.

Assembling Galaxies

The test of any scientific understanding is whether you can put all the evidence and theory together to tell the history of the objects studied. Can you describe the origin and evolution of the galaxies? Just a few decades ago, it would have been impossible, but the evidence from space telescopes and new-generation telescopes on Earth, combined with advances in computer modeling and theory, allows astronomers to outline the story of the galaxies.

Before you begin, you should eliminate a few older ideas immediately. It is so easy to imagine that galaxies normally evolve from one type to another that you could call that a **Common Misconception.** But an elliptical galaxy cannot become a spiral galaxy or an irregular galaxy because ellipticals contain almost no gas and dust from which to make new stars. That means elliptical galaxies can't be young galaxies. But you can also argue that spiral and irregular galaxies cannot evolve into elliptical galaxies because spiral and irregular galaxies contain old stars as well as young stars. The old stars tell you that spiral and irregular galaxies can't be young. The galaxy classes tell you something important, but a single galaxy does not change from one class to another any more than a cat can change into a dog.

Another old idea was that each galaxy formed from a single cloud of gas that contracted and formed stars. Astronomers call such a proposal a top-down theory. The evidence is quite clear that a galaxy does not form by a top-down contraction, but rather by a bottom-up accumulation of smaller clouds of gas and stars, the infall of gas, and the absorption of small galaxies.

Ellipticals especially appear to be the product of galaxy collisions and mergers. They are devoid of gas and dust because it was used up in the rapid star formation triggered by the interaction or blown away by the eruption of supernovae. In fact, astronomers see many **starburst galaxies** that are very luminous in the infrared because a recent collision has triggered a burst of star formation that is heating the dust (■ Figure 18-10), which reradiates the energy in the infrared. Some of these galaxies are a hundred times more luminous than our Milky Way Galaxy but so deeply shrouded in dust that they are very dim at visible wavelengths. Many such galaxies show evidence of tidal tails and are probably the result of the merger of three or more galaxies that triggered firestorms of star formation and generated tremendous clouds of dust. The Antennae Galaxies (page 411) contain over 15 billion solar masses of hydrogen gas and will become a starburst galaxy as the ongoing merger continues to trigger rapid star formation. As these galaxies use up the last of their gas and dust making stars, they will probably become normal elliptical galaxies or merge to form a single elliptical.

In contrast, spirals have never suffered major collisions. Their thin disks are delicate and would be destroyed by the tidal forces generated during a collision with a massive galaxy. Also, they retain plenty of gas and dust and continue making stars, so they have never experienced a major starburst triggered by a merger.

Of course, spiral galaxies can safely cannibalize smaller galaxies with no ill effects. The small galaxies do not generate strong enough tides to disrupt the disk of a full-sized galaxy. You have seen plenty of evidence of cannibalism in our own galaxy, and some astronomers suspect that much of the halo consists of the remains of cannibalized galaxies. Evidently our Milky Way Galaxy has not collided with a galaxy of size comparable to itself, so far. However, other evidence indicates that our galaxy will merge with the approaching Andromeda Galaxy in a few billion years, and the final result will probably be a large elliptical galaxy.

Barred spiral galaxies may also be the products of tidal interactions. Mathematical models show that bars are not stable and eventually dissipate. Tidal interactions with other galaxies may regenerate the bars. Well over half of all spiral galaxies have bars, suggesting that these tidal interactions are common.

Other processes can alter galaxies. The S0 galaxies may have lost their gas and dust as they orbited through the gas in their respective clusters. A galaxy moving rapidly through the thin gas filling a cluster would encounter a tremendous wind that could blow the galaxy's gas and dust away. In this way, a disk-shaped spiral galaxy could be reduced to an S0 galaxy. For example, X-ray observations show that the Coma cluster contains thin, hot gas between the galaxies, and astronomers have located, in a similar cluster, a galaxy in the act of plunging through such gas and being stripped of its gas and dust.

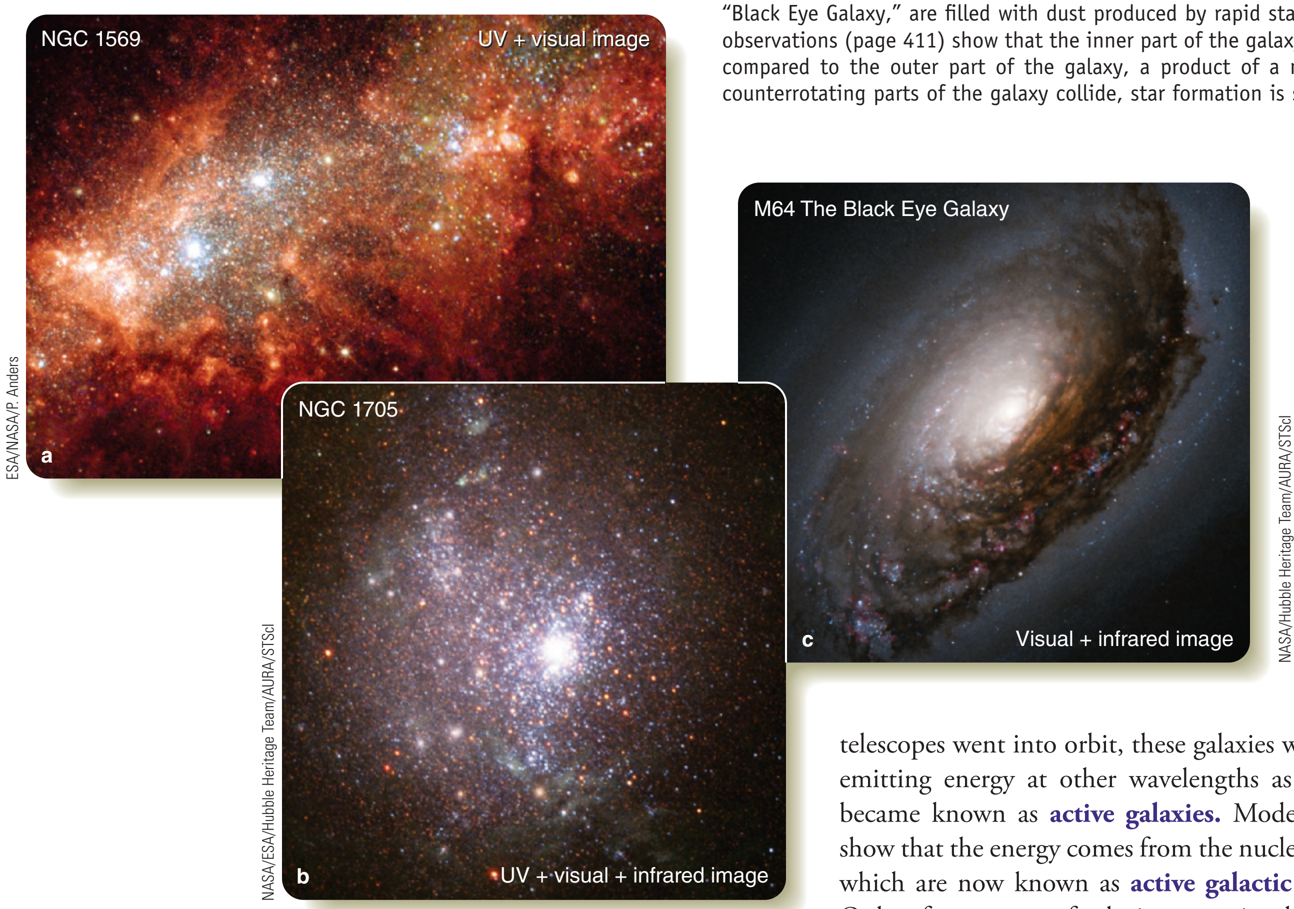


Figure 18-10

Rapid star formation: (a) NGC 1569 is a starburst galaxy filled with clouds of young stars and supernovae. At least some starbursts are triggered by interactions between galaxies. (b) The dwarf irregular galaxy NGC 1705 began a burst of star formation about 25 million years ago. (c) The inner parts of M64, known as the "Black Eye Galaxy," are filled with dust produced by rapid star formation. Radio observations (page 411) show that the inner part of the galaxy rotates backward compared to the outer part of the galaxy, a product of a merger. Where the counterrotating parts of the galaxy collide, star formation is stimulated.

Small galaxies may be produced in a number of ways. The dwarf ellipticals are too small to be produced by mergers, but they could be fragments of galaxies ripped free during interactions. Other small ellipticals may be the cores of larger galaxies that had their outer stars stripped away during encounters. In contrast, the irregular galaxies may be small fragments splashed from larger galaxies during collisions but retaining enough gas and dust to continue forming stars.

A good theory helps you understand how nature works, and astronomers are beginning to understand the exciting and complex story of the galaxies. It is already clear that galaxy evolution bears some resemblance to a pie-throwing contest.

18-4 Active Galactic Nuclei

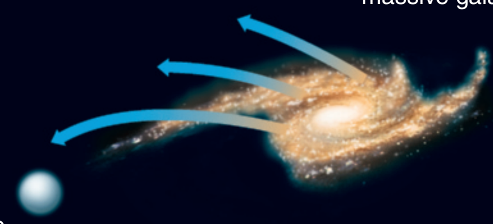
With the construction of the first large radio telescopes in the 1950s, astronomers discovered that some galaxies, dubbed **radio galaxies,** were bright at radio wavelengths. Later, when telescopes went into orbit, these galaxies were found to be emitting energy at other wavelengths as well, and they became known as **active galaxies.** Modern observations show that the energy comes from the nuclei of the galaxies, which are now known as **active galactic nuclei (AGN).** Only a few percent of galaxies are active, but they are very peculiar galaxies.

Seyfert Galaxies

In 1943, Mount Wilson astronomer Carl Seyfert published a study of spiral galaxies. Observing at visual wavelengths, Seyfert found that some spiral galaxies have small, highly luminous nuclei with peculiar spectra (■ Figure 18-11). These galaxies are now known as **Seyfert galaxies.** About 2 percent of spiral galaxies appear to be Seyfert galaxies.

The spectrum of a galaxy is the blended spectra of billions of stars, and consequently weak spectral lines get washed out. Galaxy spectra contain only a few of the strongest absorption lines found in stellar spectra. But spectra of Seyfert galaxy nuclei contain broad emission lines. Emission lines are produced by a hot, low-density gas, so the gas in the nuclei of Seyfert galaxies must be highly excited. The width of the spectral lines suggests large Doppler shifts produced by high velocities in the nuclei; gas approaching Earth produces blueshifted spectral lines, and gas going away produces redshifted lines. The combined light of gas approaching and receding, therefore, contains broad spectral

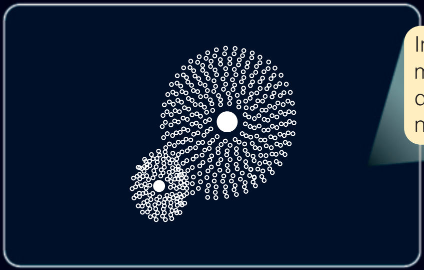
1 When two galaxies collide, they can pass through each other without stars colliding because the stars are so far apart relative to their sizes. Gas clouds and magnetic fields do collide, but the biggest effects may be tidal. Even when two galaxies just pass near each other, tides can cause dramatic effects, such as long streamers called **tidal tails.** In some cases, two galaxies can merge and form a single galaxy.



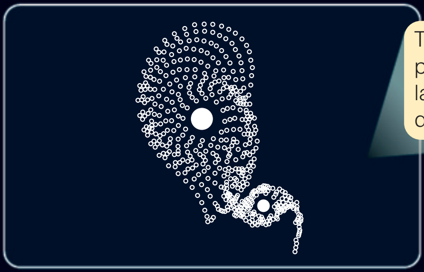
Courtesy Nicolas Martin and Rodrigo Ibata, Observatoire de Strasbourg-France

1a When a galaxy swings past a massive object such as another galaxy, tides are severe. Stars near the massive object try to move in smaller, faster orbits while stars further from the massive object follow larger, slower orbits. Such tides can distort a galaxy or even rip it apart.

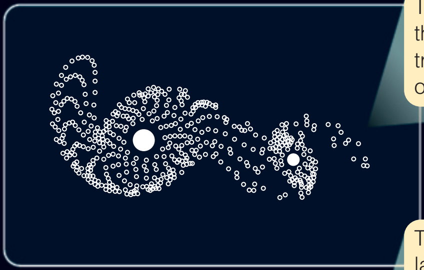
Galaxy interactions can stimulate the formation of spiral arms



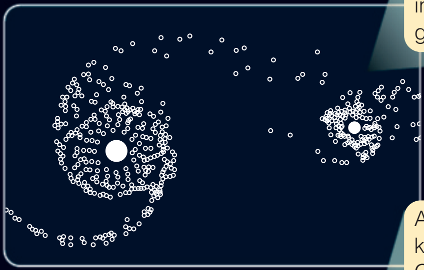
In this computer model, two uniform disk galaxies pass near each other.



The small galaxy passes behind the larger galaxy so they do not actually collide.



Tidal forces deform the galaxies and trigger the formation of spiral arms.

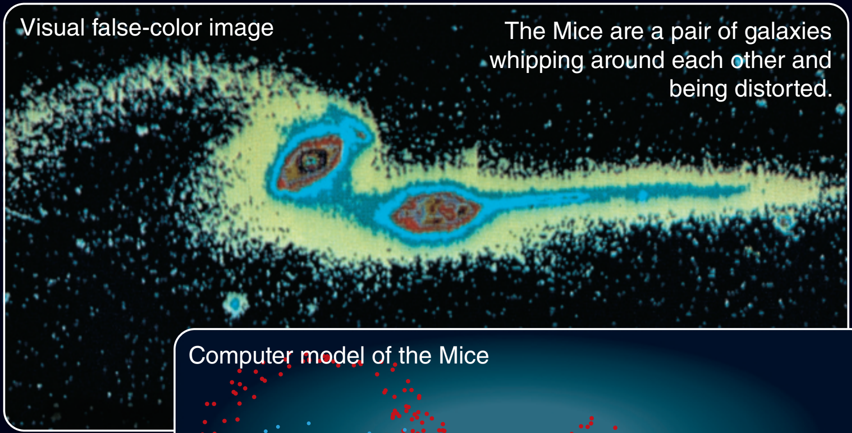


The upper arm of the large galaxy passes in front of the small galaxy.

Allen Beechel

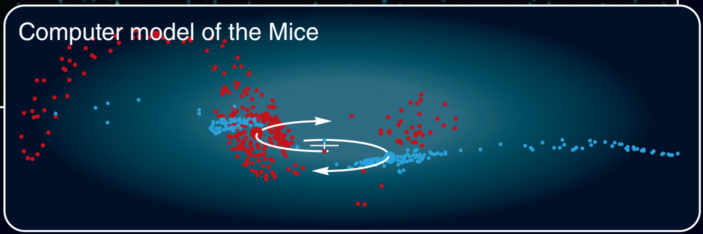
A photo of the well-known Whirlpool Galaxy resembles the computer model.

NOAO/AURA/NSF



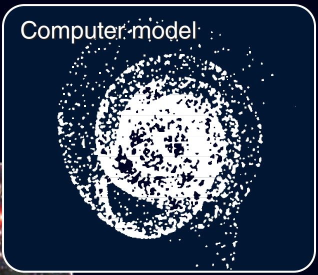
The Mice are a pair of galaxies whipping around each other and being distorted.

NOAO/AURA/NSF

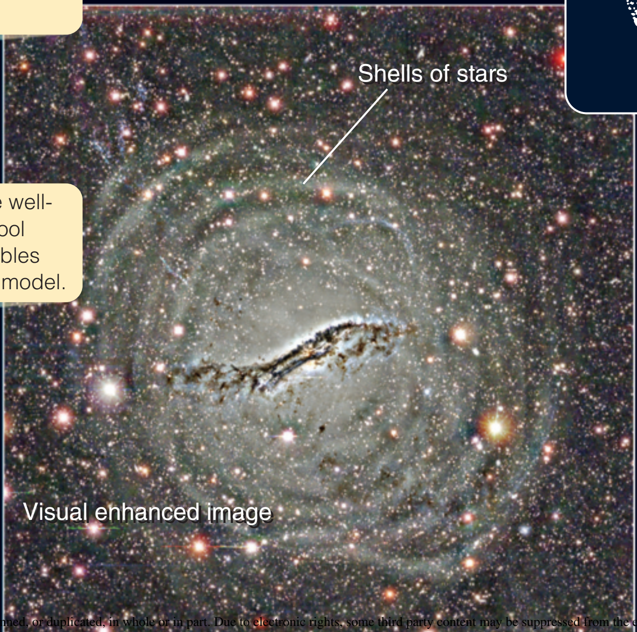


Allen Beechel

1b The merger of galaxies is called **galactic cannibalism.** Models show that merging galaxies spiral around their common center of mass while tides rip stars away and form shells.



François Schweizer and Alan Toomre

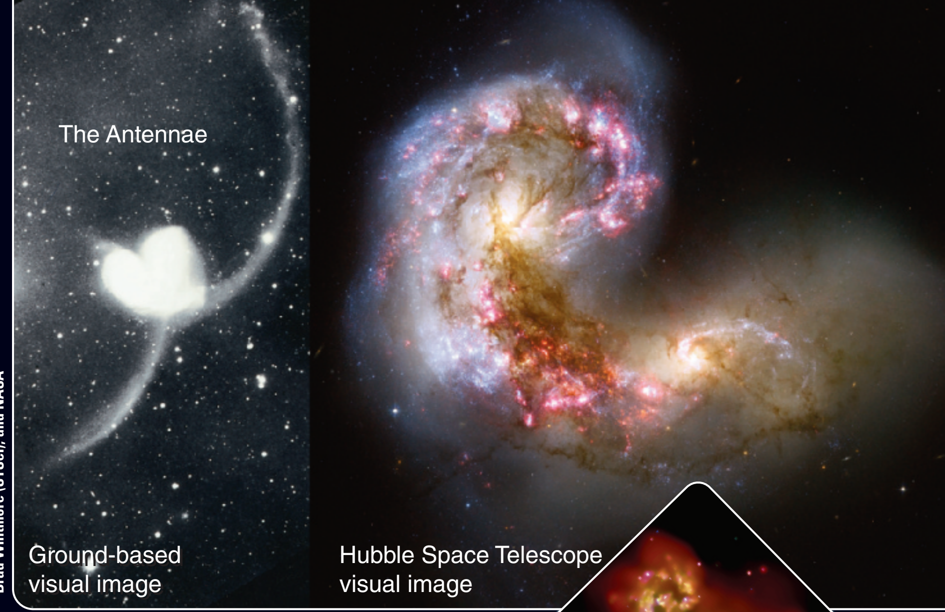


Eric Peng (JHU), Holland Ford (JHU/STScI), Ken Freeman (ANU), Rick White (STScI), NOAO/AURA/NSF

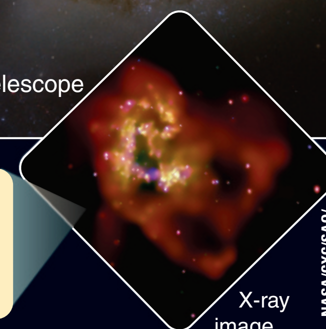
Such shells have been found around elliptical galaxies such as NGC 5128. It is peculiar in many ways and even has a belt of dusty gas. The shells revealed in this enhanced image are evidence that the giant galaxy has cannibalized at least one smaller galaxy. The giant galaxy itself may be the result of the merger of two large galaxies.

2 The collision of two galaxies can trigger firestorms of star formation as gas clouds are compressed. Galaxies NGC 4038 and 4039 have been known for years as the Antennae because the long tails visible in Earth-based photos resemble the antennae of an insect. Hubble Space Telescope images reveal that the two galaxies are blazing with star formation. Roughly a thousand massive star clusters have been born.

Spectra show that the Antennae galaxies are 10 to 20 times richer in elements such as magnesium and silicon than the Milky Way. Such metals are produced by massive stars and spread by supernova explosions.



Brad Whitmore (STScI), and NASA



An X-ray image of the Antennae shows clouds of very hot gas heated by supernovae exploding 30 times more often than in our own galaxy.

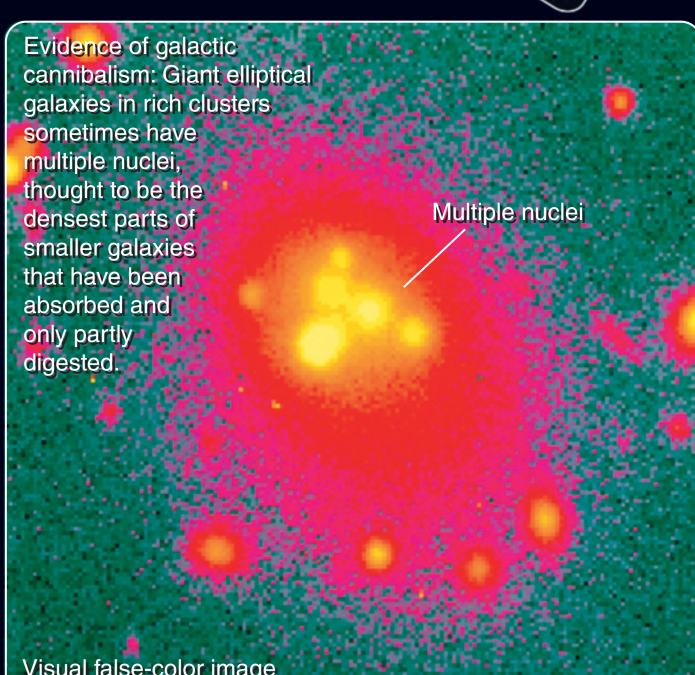
NASA/CXC/SAO/ G. Fabbiano et al.

3 Evidence of past galaxy mergers shows up in the motions inside some galaxies. NGC 7251 is a highly distorted galaxy with tidal tails in this ground-based image.

Brad Whitmore (STScI), and NASA; NASA/Heritage Hubble Team

This Hubble Space Telescope image of the core of the galaxy reveals a small spiral spinning backward in the heart of the larger galaxy.

This counter rotation suggests that NGC is the remains of two oppositely rotating galaxies that merged about a billion years ago.

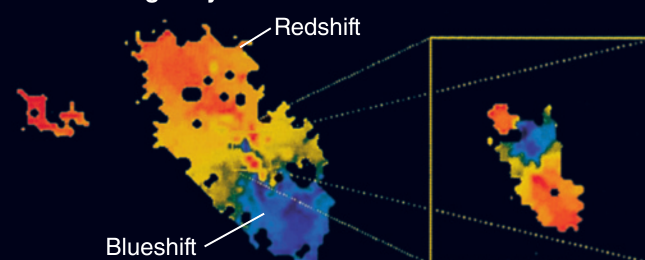


Evidence of galactic cannibalism: Giant elliptical galaxies in rich clusters sometimes have multiple nuclei, thought to be the densest parts of smaller galaxies that have been absorbed and only partly digested.

Michael J. West

3a Radio evidence of past mergers: Doppler shifts reveal the rotation of the spiral galaxy M64. The upper part of the galaxy has a redshift and is moving away from Earth, and the bottom part of the galaxy has a blueshift and is approaching. A radio map of the core of the galaxy reveals that it is rotating backward. This suggests a merger long ago between two galaxies that rotate in opposite directions.

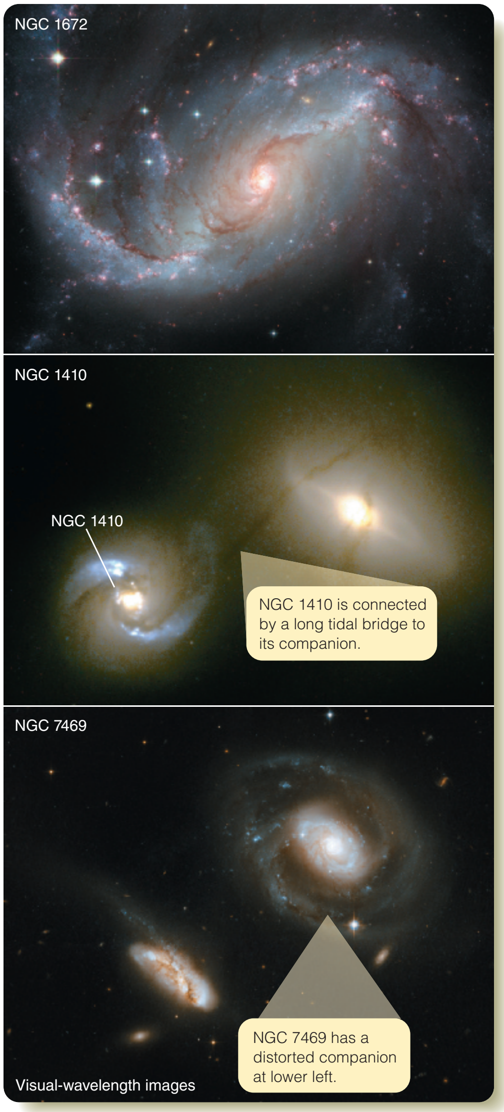
Rotation of galaxy M64



Robin Braun, NRAO/AUI/NSF

4 The Cartwheel galaxy below was once a normal galaxy but is now a **ring galaxy.** One of its smaller companions has plunged through at high speed almost perpendicular at the Cartwheel's disk. That has triggered a wave of star formation, and the more massive stars have exploded leaving behind black holes and neutron stars. Some of those are in X-ray binaries, and that makes the outer ring bright in X-rays.

NASA/JPLCaltechP. Appleton et al.; NASA/CXC/A. Wolterr & G. Trinchieri



NASA, ESA, Hubble Heritage Team, STScI/AURA-ESA/Hubble Collaboration

Wm. Keel, Univ. of Alabama

NASA, ESA, the Hubble Heritage Team, STScI/AURA-ESA/Hubble Collaboration and A. Evans, Univ. of Virginia, NRAO/Stony Brook

Figure 18-11

Seyfert galaxies are spiral galaxies with small, highly luminous nuclei. Some are interacting with nearby companions and appear distorted with tidal tails and bridges.

lines showing that the velocities at the center of Seyfert galaxies are roughly 10,000 km/s, about 30 times greater than velocities at the center of normal galaxies. Something violent is happening in the cores of Seyfert galaxies.

The brilliant nuclei of Seyfert galaxies fluctuate rapidly, especially at X-ray wavelengths. A Seyfert nucleus can change its X-ray brightness by a significant amount in only minutes. As you learned when you studied neutron stars, an object cannot change its brightness in a time shorter than the time it takes light to cross its diameter. If the Seyfert nucleus can change in a few minutes, then it cannot be more than a few light-minutes in diameter—about the size of Earth's orbit. In spite of their small size, the cores of Seyfert galaxies produce tremendous amounts of energy. The brightest emit a hundred times more energy than the entire Milky Way Galaxy.

Lots of energy is produced in a very small volume with extremely high temperatures and velocities. What does that remind you of? Astronomers conclude that the centers of these galaxies contain supermassive black holes orbited by hot accretion disks through which matter is flowing into the black hole.

Earlier in this chapter, you saw evidence that most galaxies contain supermassive black holes at their centers. But most galaxies are not erupting. The shapes of Seyfert galaxies provide an important clue. About 25 percent of Seyfert galaxies have peculiar shapes suggesting that they have had tidal interactions with other galaxies (■ Figure 18-12). There is also statistical evidence (**How Do We Know? 18-3**) that Seyfert galaxies are more common in interacting pairs of galaxies than in isolated galaxies. These clues hint that Seyfert galaxies may have been triggered into activity by collisions or interactions with companions. You will find more such evidence as you study other kinds of active galaxies.

Double-Lobed Radio Sources

Beginning in the 1950s, radio astronomers found that some sources of radio energy in the sky consisted of pairs of radio-bright regions. When optical telescopes studied the locations of these radio sources, they revealed galaxies located between the two regions emitting radio energy, and the galaxies were dubbed **double-lobed radio galaxies.** Unlike Seyfert galaxies, which emit intense radiation from their cores, these radio galaxies were producing energy from two external radio lobes.

Study **Cosmic Jets and Radio Lobes** on pages 414–415 and notice four important points and two new terms:

1. The shapes of radio lobes suggest that they are inflated by jets of excited gas emerging from the nucleus of the central galaxy. This has been called the *double-exhaust model.* The presence of *hot spots* and synchrotron radiation shows that the jets are very powerful.
2. Active galaxies that have jets and radio lobes are often deformed or interacting with other galaxies.
3. The complex shapes of some jets and radio lobes can be explained by the motions of the active galactic nuclei. A good example of this is 3C 31 (the 31st source in the *Third Cambridge Catalog of Radio Sources*) with its twisting radio lobes.
4. These jets are consistent with matter falling into a central supermassive black hole. You have seen similar jets produced by accretion disks around protostars, neutron stars, and stellar mass black holes; although the details are not entirely understood, the same process seems to be producing all of these jets.

How Do We Know? | 18-3

Statistical Evidence

How can statistics be useful if they can't be specific? Some scientific evidence is statistical. Observations suggest, for example, that Seyfert galaxies are more likely to be interacting with a nearby companion than a normal galaxy is. This is statistical evidence, so you can't be certain that any specific Seyfert galaxy will have a companion. How can scientists use statistical evidence to learn about nature when statistics contain built-in uncertainty?

Meteorologists use statistics to determine how frequently storms of a certain size are likely to occur. Small storms happen every year, but medium-sized storms may happen on average only every ten years. Hundred-year storms are much more powerful but occur much less frequently—on average only once in a hundred years.

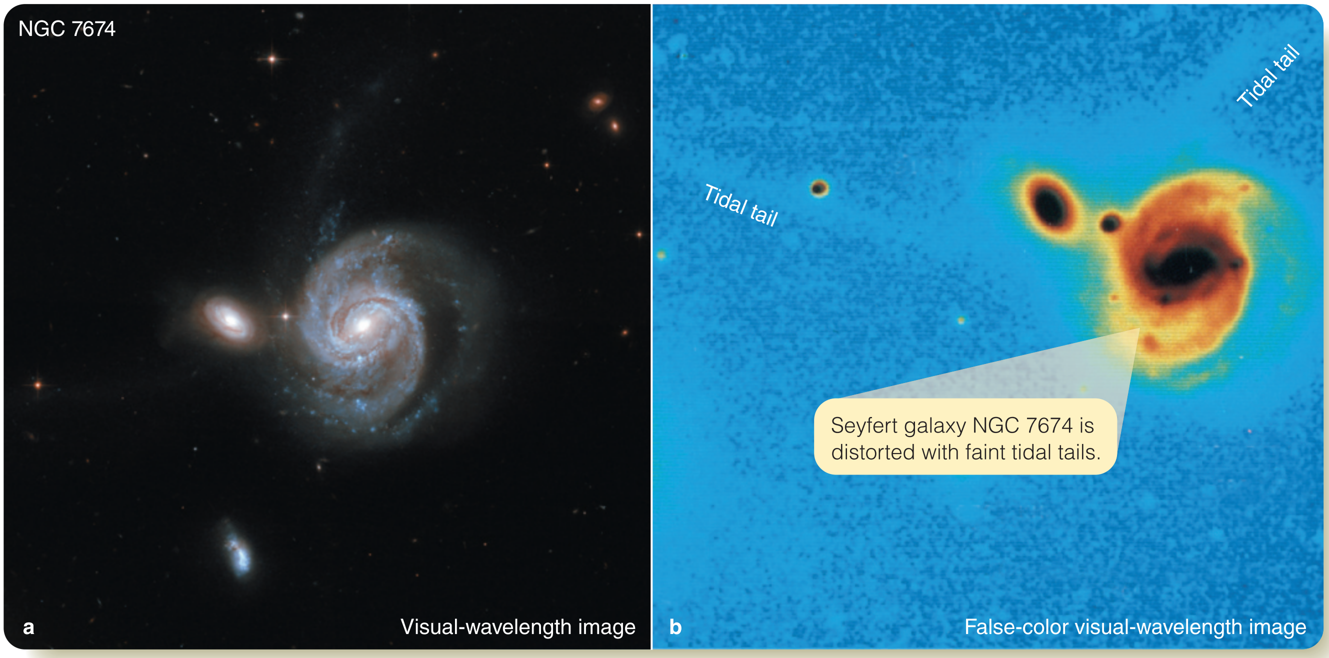
Those meteorological statistics can help you make informed decisions—as long as you understand the powers and limitations of statistics. Would you buy a house protected from a river by a levee that was not designed to withstand a hundred-year storm? In any one year, the chance of your house being destroyed would be only 1 in 100. You know the storm will hit eventually, but you don't know when. If you buy the house, a storm might destroy the levee the next year, but you might own the house for your whole life and never see a hundred-year storm. The statistics can't tell you anything about a specific year.

Before you buy that house, there is an important question you should ask the meteorologists. "How much data do you have on storms?" If they have only ten years of data, then they don't really know much about hundred-year storms. If they have three centuries of data, then their statistical data are significant.

Sometimes people dismiss important warnings by saying, "Oh, that's only statistics." Scientists can use statistical evidence if it passes two tests. It cannot be used to draw conclusions about specific cases, and it must be based on large enough data samples so the statistics are significant. With these restrictions, statistical evidence can be a powerful scientific tool.

Marko Georgiev/Getty Images

Statistics can tell you that a bad storm will eventually hit, but it can't tell you when.



NASA, ESA, the Hubble Heritage Team, STScI/AURA-ESA/Hubble Collaboration and A. Evans, Univ. of Virginia, NRAO/Stony Brook

John W. Mackenty, Institute for Astronomy, University of Hawai'i

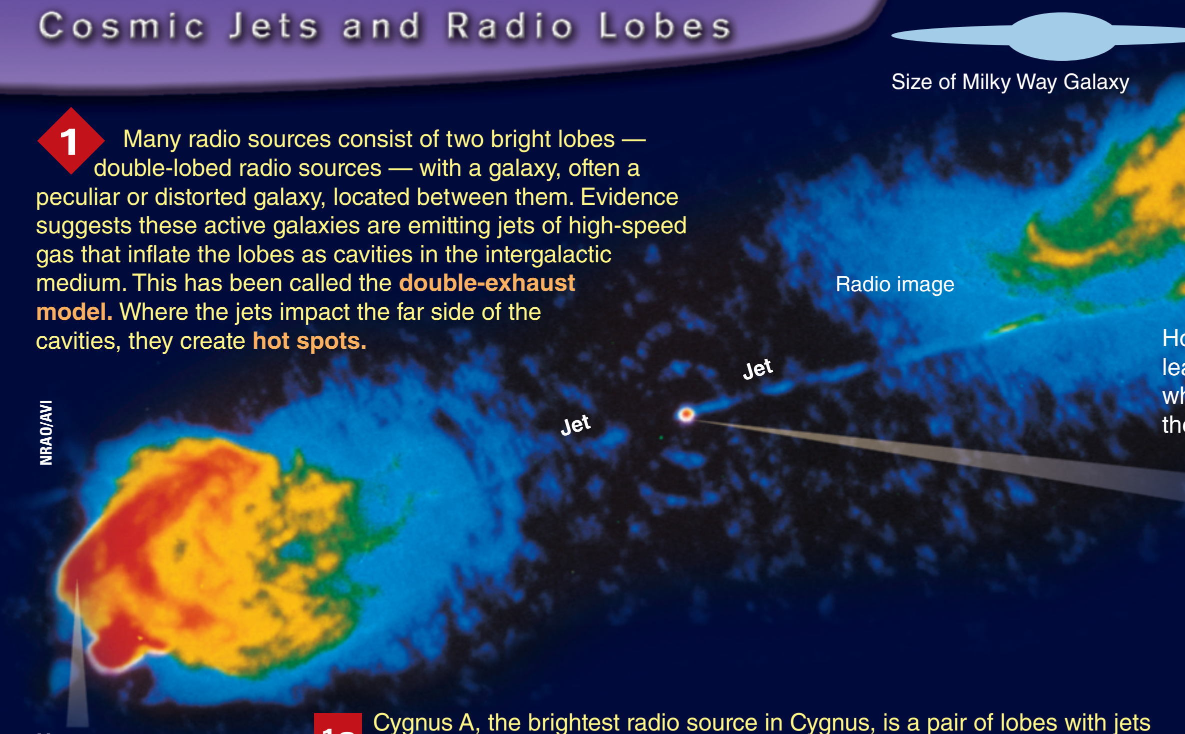
© Cengage Learning 2014

Figure 18-12

(a) The Seyfert galaxy NGC 7674 is part of a small, compact group of galaxies and is interacting with its smaller companion. Tidal tails are visible extending to the left and to the upper right. (b) The tidal tails are more easily visible in this enhanced image.

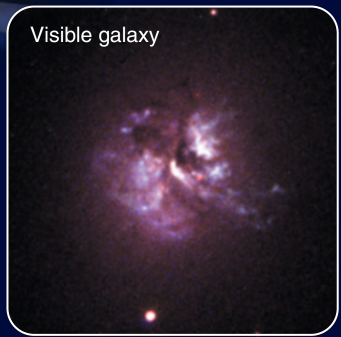
Cosmic Jets and Radio Lobes

1 Many radio sources consist of two bright lobes — double-lobed radio sources — with a galaxy, often a peculiar or distorted galaxy, located between them. Evidence suggests these active galaxies are emitting jets of high-speed gas that inflate the lobes as cavities in the intergalactic medium. This has been called the **double-exhaust model.** Where the jets impact the far side of the cavities, they create **hot spots.**



Hot spots lie on the leading edge of a lobe where the jet pushes into the surrounding gas.

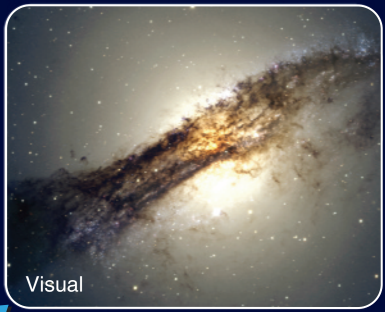
NRAO/AVI



Bob Fosbury/European Southern Observatory

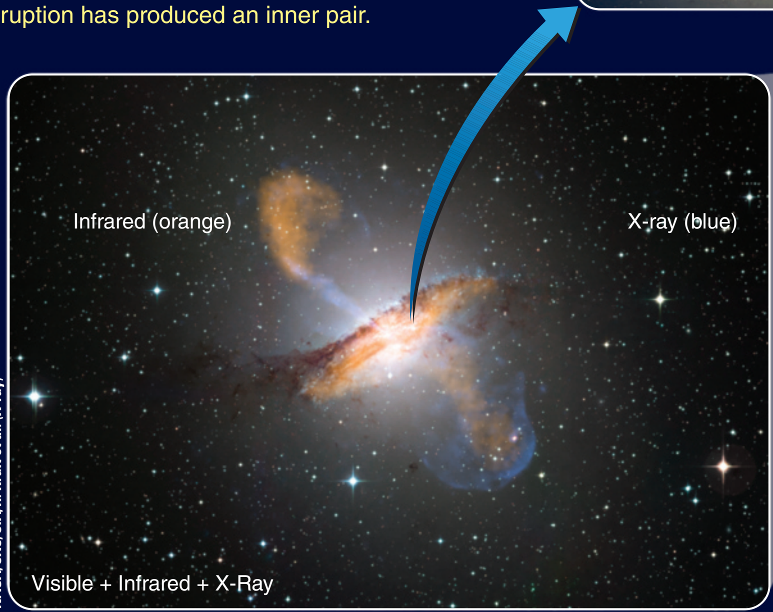
1a Cygnus A, the brightest radio source in Cygnus, is a pair of lobes with jets leading from the nucleus of a highly disturbed galaxy. In this false-color image, the areas of strongest radio signals are shown in red and the weakest in blue. Because the radio energy detected is synchrotron radiation, astronomers conclude that the jets and lobes contain very-high-speed electrons, usually called relativistic electrons, spiraling through magnetic fields about 1000 times weaker than Earth's field. The total energy in a radio lobe is about 10^{53} J — what you would get if you turned the mass of a million suns directly into energy.

2 Radio galaxy NGC 5128 lies between two radio lobes, and, like many active galaxies, is strangely distorted. The dust ring rotates about an axis perpendicular to the ring, but the spherical cloud of stars rotates about an axis that lies in the plane of the ring. NGC 5128 appears to be two galaxies, a giant elliptical and a spiral, passing through each other. This has triggered multiple eruptions. An earlier eruption has produced a large outer pair of lobes, and a more recent eruption has produced an inner pair.



ESO

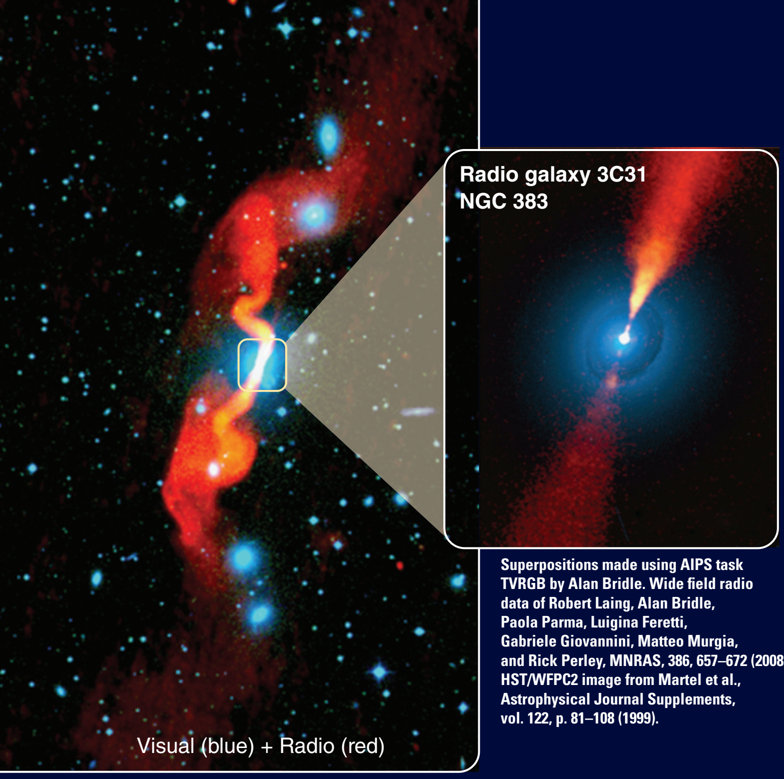
2a At visual wavelengths, NGC 5128 (left) appears to be a giant elliptical galaxy with a strange belt of dust. It is known to radio astronomers as Centaurus A. Detailed observations at many wavelengths reveal that the galaxy's core contains a supermassive black hole of about 100 million solar masses ejecting jets in opposite directions. The jets are visible in the image at lower left.



ESO/WFI (Optical); MPIfR/ESO/APEX/A. Weiss et al. (Submillimetre); NASA/CXC/CfA/R. Kraft et al. (X-ray)

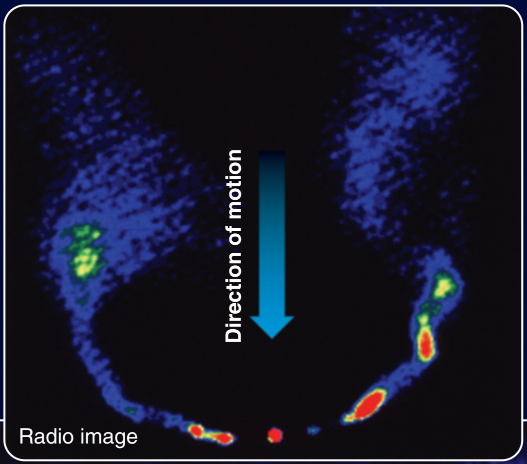
If the outer radio lobes of Centaurus A were visible to your eyes, they would look 10 times larger than the full moon.

Line art on this page © Cengage Learning 2014



Superpositions made using AIPS task TVRGB by Alan Bridle. Wide field radio data of Robert Laing, Alan Bridle, Paola Parma, Luigina Feretti, Gabriele Giovannini, Matteo Murgia, and Rick Perley, MNRAS, 386, 657–672 (2008). HST/WFPC2 image from Martel et al., Astrophysical Journal Supplements, vol. 122, p. 81–108 (1999).

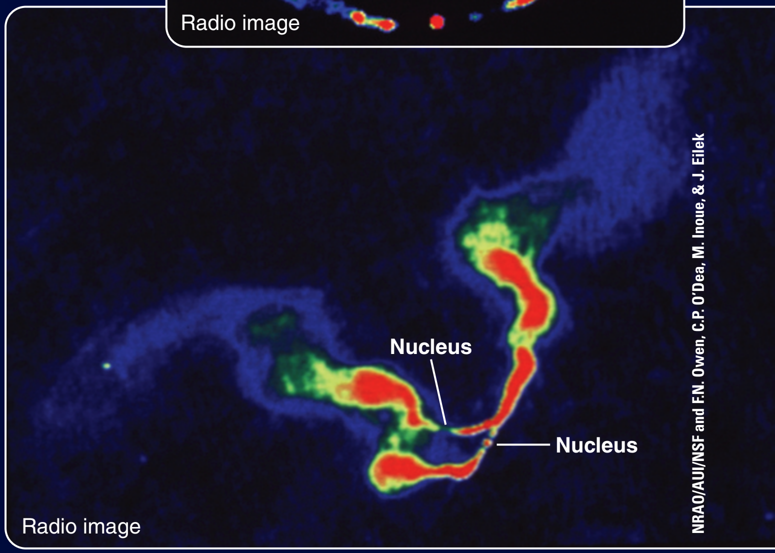
3 The radio jets from NGC 1265 (right) are being left behind as the galaxy moves rapidly through the gas of the intergalactic medium. Twists in the tails are presumably caused by motions of the active nucleus.



NRAO/AUI/NSF and C. O'Dea & F. Owen

3a Radio source 3C31 (above) is one galaxy in a chain of galaxies. It is ejecting jets from its core in opposite directions, and the jets appear to twist and turn because the active nucleus is orbiting another object such as the nucleus of a recently absorbed galaxy.

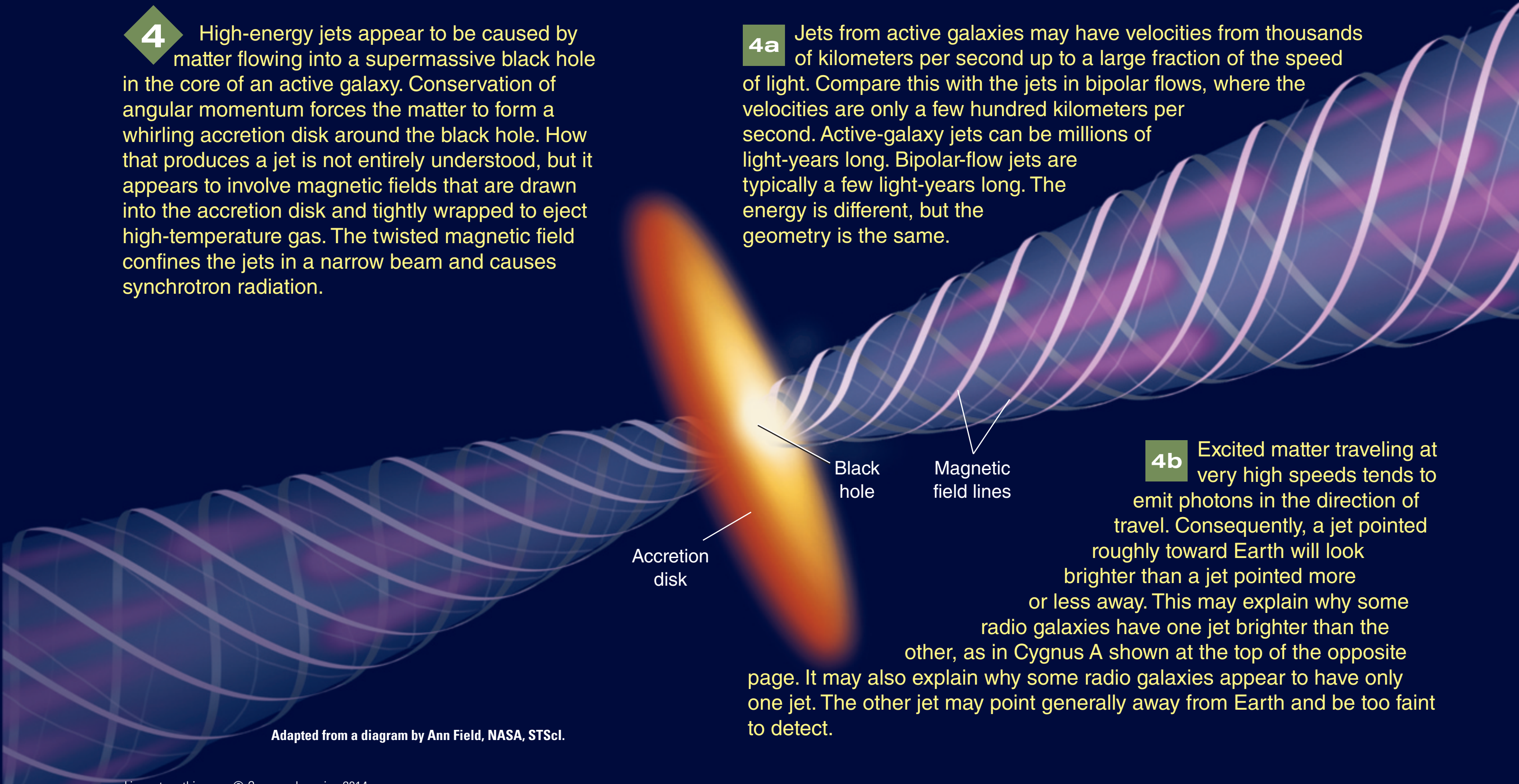
3b The radio source 3C75 (right) is produced by two galaxies experiencing a close interaction. They may be merging. As the active nuclei whip around each other, their jets appear to twist and turn. The size of the visible galaxies in this image would be about the size of cherries.



NRAO/AUI/NSF and F.N. Owen, C.P. O'Dea, M. Inoue, & J. Eilek

4 High-energy jets appear to be caused by matter flowing into a supermassive black hole in the core of an active galaxy. Conservation of angular momentum forces the matter to form a whirling accretion disk around the black hole. How that produces a jet is not entirely understood, but it appears to involve magnetic fields that are drawn into the accretion disk and tightly wrapped to eject high-temperature gas. The twisted magnetic field confines the jets in a narrow beam and causes synchrotron radiation.

4a Jets from active galaxies may have velocities from thousands of kilometers per second up to a large fraction of the speed of light. Compare this with the jets in bipolar flows, where the velocities are only a few hundred kilometers per second. Active-galaxy jets can be millions of light-years long. Bipolar-flow jets are typically a few light-years long. The energy is different, but the geometry is the same.



4b Excited matter traveling at very high speeds tends to emit photons in the direction of travel. Consequently, a jet pointed roughly toward Earth will look brighter than a jet pointed more or less away. This may explain why some radio galaxies have one jet brighter than the other, as in Cygnus A shown at the top of the opposite page. It may also explain why some radio galaxies appear to have only one jet. The other jet may point generally away from Earth and be too faint to detect.

Adapted from a diagram by Ann Field, NASA, STScI.

The evidence shows that the cores of many galaxies are occupied by supermassive black holes with matter flowing inward. As the matter falls toward the center it releases tremendous gravitational energy and becomes intensely hot. The hot accretion disks can emit X-rays and eject jets in opposite directions.

Quasars

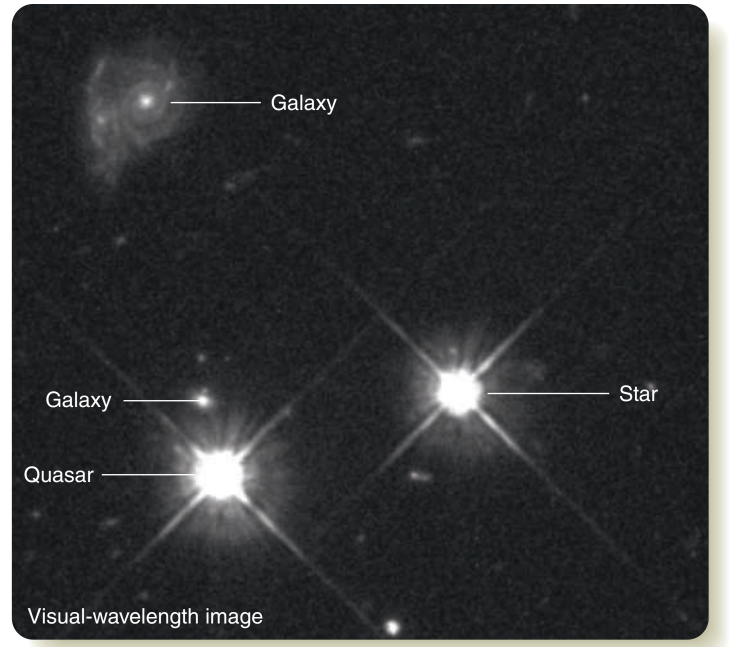
Through the 1950s, radio astronomers became familiar with celestial radio sources that were either huge clouds of gas or distant radio galaxies, so they were surprised in the early 1960s when some radio sources turned out to look like stars in visual-wavelength photographs (■ Figure 18-13). First called quasi-stellar objects, they were soon referred to as **quasars.** Many more quasars have been found over the years, and most are radio silent. The radio astronomers stumbled over those emitting radio energy because they were easy to notice. From their discovery, quasars have been a puzzle to astronomers.

The spectra of quasars were strange in that they contained a few unidentified emission lines superimposed on a continuous spectrum. In 1963, Maarten Schmidt at Hale Observatories tried redshifting the hydrogen Balmer lines to see if they could be made to agree with the lines in the spectrum of the quasar known as 3C 273. At a redshift of 15.8 percent, three lines clicked into place (■ Figure 18-14). Other quasar spectra quickly yielded to this approach, revealing even larger redshifts.

Numerically the redshift z is the change in wavelength $\Delta\lambda$ divided by the unshifted wavelength λ_o:

$$\text{redshift} = \frac{\Delta\lambda}{\lambda_o}$$

According to the Hubble law, these large redshifts imply very large distances. The first quasars studied were the brighter ones, but surveys have found many more. For example, the Sloan Digital Sky Survey discovered 90,000. Most of these quasars have very high redshifts and lie very far from Earth.

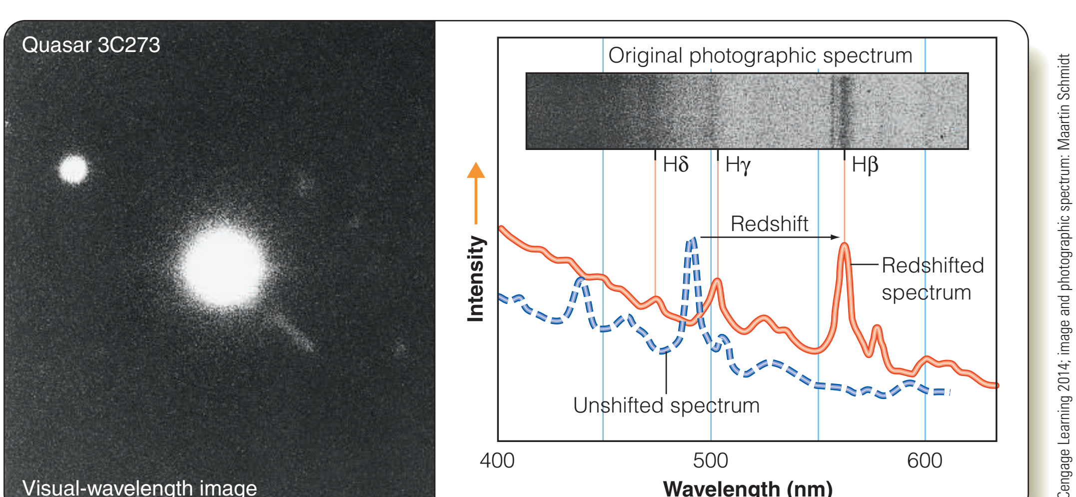


■ **Figure 18-13**

Quasars have starlike images clearly different from the images of even very distant galaxies. Although quasars look like stars, their spectra are unlike the spectra of stars or galaxies. The spikes on these images were produced by diffraction in the telescope.

Many quasar redshifts are greater than 1.0, and that may strike you as impossible. The Doppler formula implies that such objects must have velocities greater than the speed of light. But the redshifts of the galaxies and quasars are not produced by the Doppler effect. As you will discover in the next chapter, the redshifts are produced by the expansion of the universe, and astronomers must use the equations of general relativity to interpret them. Redshifts greater than 1.0 are not a problem and merely indicate great distance.

Although the quasars are far away, they are not very faint. A typical galaxy at such a distance would be faint and extremely difficult to detect, but quasars show



■ **Figure 18-14**

This image of 3C 273 shows the bright quasar at the center surrounded by faint fuzz. Note the jet protruding to lower right. The original photographic plate holding a spectrum of 3C 273 contains three hydrogen Balmer lines, H_δ, H_γ, and H_β. The spectrum is redshifted by 15.8 percent. The dashed line shows the unshifted position of the spectrum.

up on photographs as noticeable points of light. If you put the apparent brightnesses and huge distances of quasars into the magnitude–distance relation (See Reasoning with Numbers 13-2), you will discover that quasars are ultraluminous, having 10 to 1000 times the luminosity of a large galaxy.

Soon after quasars were discovered, astronomers detected fluctuations in brightness over times as short as a few hours. Those rapid fluctuations show that quasars are small objects, not more than a few light-hours in diameter—smaller than our solar system.

By the late 1960s astronomers trying to understand quasars faced a problem: How could quasars be ultraluminous but also very small? What could make 10 to 1000 times more energy than a galaxy in a region as small as our solar system? Since that time, new, large telescopes in space and on Earth's surface have revealed that quasars are often surrounded by hazy features whose spectra resemble those of normal galaxies. Evidently, quasars are located in galaxies. In addition, radio telescopes have revealed that some quasars are ejecting jets and inflating radio lobes. The evidence is now overwhelming that quasars are the active cores of very distant galaxies. In other words, quasars are the most extreme kind of active galactic nuclei.

SCIENTIFIC ARGUMENT

What evidence suggests that radio lobes are inflated by jets from the nucleus?

Of course, the strongest evidence is that in a few cases a jet is actually detectable at visual or radio wavelengths leading from the center of a galaxy out into a radio lobe. But astronomers also note that hot spots occur on the outer edges of many radio lobes where a jet, whether it is visible or not, would collide with the intergalactic medium. In some cases where jets are detectable, they are curved and twisted by the orbital motion of the nucleus, which shows that they are being produced by gas expelled from the nucleus.

Evidence is the key to understanding science. Now create a new scientific argument to answer the following: **What evidence suggests that the cores of active galaxies produce lots of energy in a small region?**

18-5 Supermassive Black Holes

You now know that many galaxies contain supermassive black holes at their centers. Why do such objects produce eruptions, and how did they form?

Disks and Jets

Matter flowing inward toward a black hole spins very fast and becomes very hot. It spins because it must conserve angular momentum as it sinks inward where it forms a flattened disk around the central black hole. Supermassive black holes have stronger gravity than stellar-mass black holes and produce faster spins for the infalling matter and higher temperatures. Even a supermassive black hole is surprisingly small. A 10-million-solar-mass black hole would be only one fifth the diameter of Earth's orbit, so the matter can get close to the black hole and orbit very fast. The infalling matter heats up because it picks up speed, and as it collides with other matter, those high velocities become thermal energy. That is, matter falling inward converts gravitational energy into thermal energy and becomes hot.

Theoretical calculations predict that that the high temperature "puffs up" the inner part of the disk and makes it thick. Even closer to the black hole, an orbiting particle is unstable and must spiral into the black hole, so the innermost part of the disk is empty and the black hole is hidden deep inside an empty central well. Further out, the disk is thinner and cooler, but the outermost part of the disk, according to calculations, is a fat, cool torus (doughnut shape) of dusty gas.

Astronomers can't see black holes, but in some active galaxies the Hubble Space Telescope can detect the outer parts of the central disks (■ Figure 18-15). Spectra reveal the speed of rotation, and Kepler's third law yields the mass of the central object. Some supermassive black holes have masses of a few million solar masses, like the one in the center of our Milky Way Galaxy, but the most massive contain billions of solar masses.

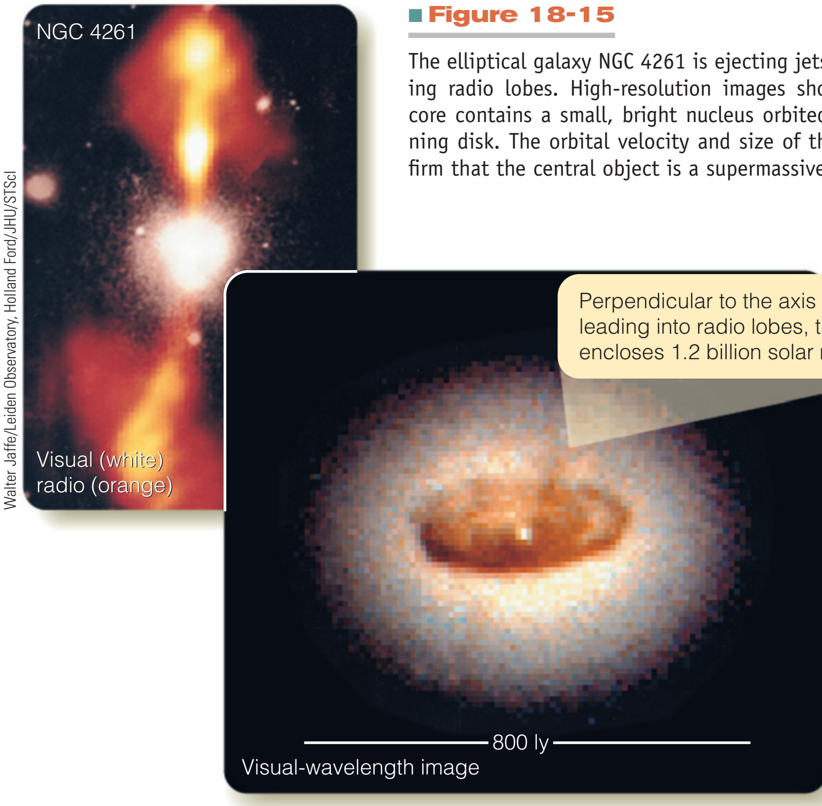
The inner part of the accretion disk around a supermassive black hole can reach temperatures of millions of degrees and emit X-rays. In fact, cosmic ray detectors on Earth have observed ultra-high-energy particles that come from places in the sky occupied by active galactic nuclei. These powerful supermassive black holes and their hot accretion disks may be filling the universe with high-speed particles packing the wallop of major league fastballs.

No one knows exactly how the supermassive black hole and its disk produce jets of gas and radiation, but magnetic fields are a factor. Because the disk is at least partially ionized, magnetic fields are trapped in the gas of the disk, drawn inward, and wound up. Theorists suggest that this creates powerful magnetic tubes extending along the axis of rotation, channeling hot gas outward in opposite directions. The jets seem to originate very close to the supermassive black hole and are then focused and confined by the enclosing magnetic tubes.

The mechanism that produces jets is understood in only a general way, but astronomers are now trying to work out the details. How can supermassive black holes explain all of the different kinds of active galaxies that are observed?

The Search for a Unified Model

When a field of research is young, scientists find many seemingly different phenomena, such as Seyfert galaxies, double-lobed radio galaxies, quasars, cosmic jets, and so on. As the



Walter Jaffe/Leiden Observatory, Holland Ford/JHU/STScI

L. Ferrarese (Johns Hopkins University) and NASA

■ **Figure 18-15**

The elliptical galaxy NGC 4261 is ejecting jets and inflating radio lobes. High-resolution images show that the core contains a small, bright nucleus orbited by a spinning disk. The orbital velocity and size of the disk confirm that the central object is a supermassive black hole.

research matures, scientists begin seeing similarities and eventually are able to unify the different phenomena as different aspects of a single process. This organization of evidence and theory into logical arguments that explain how nature works is the real goal of science. Astronomers studying active galaxies have developed a **unified model** of active galaxy cores that is well supported by evidence. A monster black hole is the centerpiece.

According to the unified model, what you see when you view the core of an active galaxy depends on how the black hole's accretion disk is tipped with respect to your line of sight (■ Figure 18-16). You should note that the accretion disk may be tipped at a steep angle to the plane of its galaxy, so just because you see a galaxy face-on doesn't mean you are looking at the accretion disk face-on.

The spectrum you observe depends on the angle of the disk. If you view the accretion disk edge-on, you cannot see into the central area at all, and you see lower temperatures and velocities. If the disk is tipped so you can see the hot, inner gas, you see higher temperatures and velocities. If the disk is nearly face-on, you would be looking directly into the central cavity—down the dragon's throat—and you would see extreme conditions. A few active galaxies have such extreme spectra.

The unified model is far from complete. The actual structure of accretion disks is poorly understood, as is the process by which the disks produce jets. The unified model does not explain all of the differences among active galaxies and quasars. Rather, it is a model that provides some clues to what is happening in active galactic nuclei.

Triggering Eruptions

Most galaxies contain supermassive black holes at their centers, but only a few percent of galaxies have active galactic nuclei. That must mean that most of the supermassive black holes are dormant. They are sleeping.

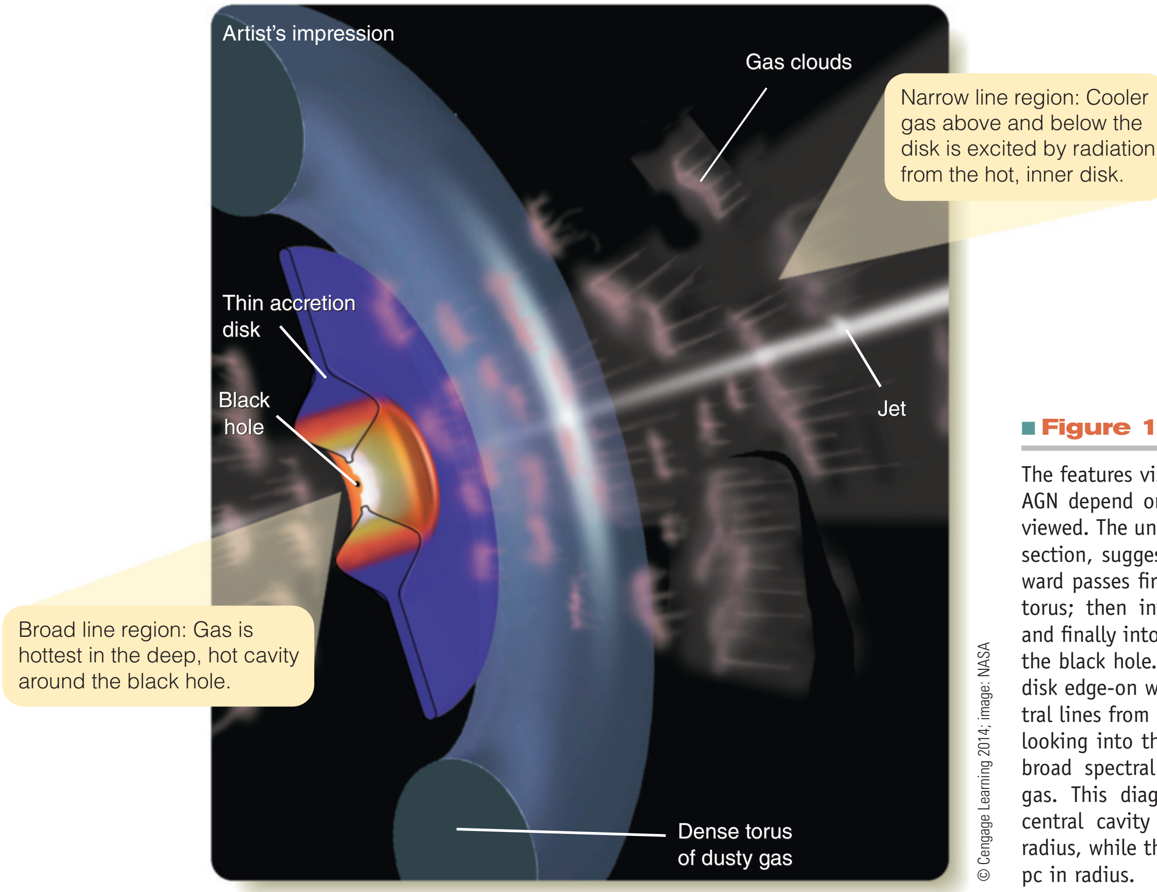
What could trigger a supermassive black hole to erupt? The answer is something that you studied back in Chapter 4—tides. Tides twist interacting galaxies and rip matter away into tidal tails, but mathematical models show that those same interactions can also throw matter inward. Even a small amount of matter falling into a black hole can produce an outburst (■ Figure 18-17) A sudden flood of matter flowing into the accretion disk around a supermassive black hole would trigger it into eruption. This explains why active galaxies are often distorted; they have been twisted by tidal forces as they interacted or merged with another galaxy. Some active galaxies have nearby companions, and you can suspect that the companions are guilty of tidally distorting the other galaxy and triggering an eruption. Images of some quasars reveal that they are embedded in galaxies that are distorted and lie near other distorted galaxies (■ Figure 18-18). Statistical evidence suggests there are other processes that can throw matter inward and trigger an eruption, but astronomers are just beginning to sort out that puzzle.

The Coevolution of Galaxies and Black Holes

Evidence shows that galaxies and supermassive black holes have evolved together. Astronomers can see this progression because at great distances the look-back time is large and observations reveal the universe as it was long ago.

The masses of supermassive black holes are related to the masses of the host galaxies' central bulges. In each case, the mass of the black hole is about 0.1 percent the mass of the surrounding central bulge. But there is no relation to the masses of the disks of the galaxies. This reveals that the formation of supermassive black holes was connected with the formation of the central bulges. They formed and evolved together.

In the next chapter, you will see evidence that the universe began 13.7 billion years ago and that it has been expanding ever since. Within a few hundred million years, the first clouds of gas began forming stars and falling together to form the great



© Cengage Learning 2014; image: NASA

Figure 18-16

The features visible in the spectrum of an AGN depend on the angle at which it is viewed. The unified model, shown in cross section, suggests that matter flowing inward passes first through a large, opaque torus; then into a thinner, hotter disk; and finally into a small, hot cavity around the black hole. Telescopes viewing such a disk edge-on would see only narrow spectral lines from cooler gas, but a telescope looking into the central cavity would see broad spectral lines formed by the hot gas. This diagram is not to scale. The central cavity may be only 0.01 pc in radius, while the outer torus may be 1000 pc in radius.

star clouds that became the nuclear bulges of galaxies. The supermassive black holes formed at the same time. In fact, some evidence suggests that the black holes formed first and pulled matter inward to form central bulges. Matter flooding into these black holes would have triggered powerful outbursts. The formation of a central bulge was evidently a violent process. Eruptions of the growing central black hole could have pushed away infalling gas and limited the material available to form the stars of the central bulge. Also, infalling matter can trigger bursts of star formation, and the resulting supernovae can blow gas out of a galaxy and also limit bulge growth.

Some of the most distant quasars could have been caused by the formation of supermassive black holes, but these extremely distant objects are difficult to image with existing telescopes. Most of the quasars and other active galaxies that astronomers see with today's telescopes have been triggered into eruption by the interaction, collision, and merging of galaxies, a process that throws matter into the central black holes.

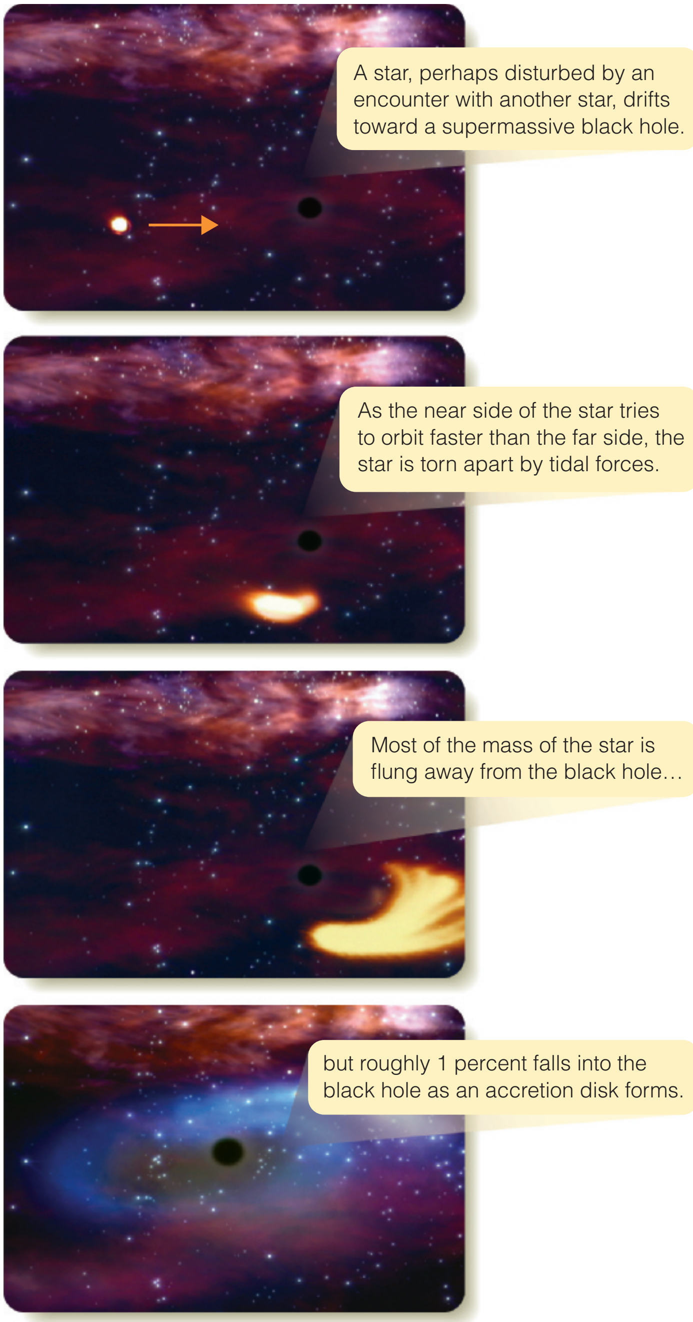
Collisions have been important in the evolution of galaxies and supermassive black holes. When the universe was young and had not expanded very much, galaxies were closer together and collided more often. Also, as small galaxies fell inward and formed the halos and disks of larger galaxies, more matter would have been fed inward to the supermassive black holes and triggered more eruptions.

Observations show that in the first half of the history of the universe each galaxy merged with other galaxies as often as three times every billion years. Most mergers were with smaller galaxies, but a few would have been with large galaxies and triggered eruptions visible from Earth today as quasars.

Quasars are most common with redshifts over 2 and less common with redshifts above 3. The largest quasar redshifts are over 6, but such high-redshift quasars are rare. Evidently astronomers see few quasars at high redshifts because they are looking back to an age when the universe was so young it had not yet formed many galaxies. At redshifts between 2 and 3, galaxies were actively growing and colliding, and quasars were about 1000 times more common than they are now. Nevertheless, even during the age of quasars, quasar eruptions must have been unusual. At any one time, only a small fraction of galaxies had quasars erupting in their cores. As smaller galaxies were gobbled up and the expansion of the universe carried the galaxies away from each other, galaxy formation became less active, interactions between galaxies became rare, and quasars became even less common.

"Then where are all the dead quasars?" an astronomer asked recently. "There is no way to get rid of supermassive black holes, so all of the galaxies that had short-lived quasars must

Star Falling into a Black Hole



© Cengage Learning 2014; image: ESA and Stephanie Komossa (Max Planck Institute for Extraterrestrial Physics)

Figure 18-17

Orbiting X-ray telescopes observing active galaxies sometimes detect X-ray flares equaling the energy of a supernova explosion. Such flares are evidently caused when a star wanders too close to the supermassive black hole at the center of the galaxy and tidal forces rip the star apart.

X-ray (NASA/CXC/SAO/P. Green et al.), Optical (Carnegie Obs./Magellan/W.Baade Telescope/J.S.Mulchaey et al.)

Figure 18-18

Two quasars (center) are embedded in merging galaxies. Tidal forces have flung out tidal tails, a sure sign of galaxies merging.

still have those supermassive black holes." Where are all those dead quasars today? You know where to look—the cores of galaxies.

Most large galaxies contain supermassive black holes at their centers, but the black holes have consumed most of the nearby gas, dust, and stars and are now dormant. Our own Milky Way Galaxy could have had a quasar at its core long ago, but now its central supermassive black hole is on a strict diet. A slow trickle of matter flowing into the black hole could explain the mild activity seen there. Occasional interactions between galaxies can throw matter inward and awaken a supermassive black hole into eruption as a Seyfert galaxy or a double-lobed radio galaxy.

The active galaxies are not a rare kind of galaxy. They are normal galaxies passing through a stage that many galaxies experience: They are not really peculiar. They are an important part of the story of the formation and evolution of galaxies.

Quasars are most common with redshifts of a little over 2 and less common with redshifts above 3. The largest quasar redshifts are over 6, and such high-redshift quasars are quite rare. Evidently astronomers see few quasars at high redshifts because they are looking back to an age when the universe was so young it had not yet formed many galaxies. At redshifts between about 2 and 3, galaxies were actively growing, colliding, and merging, and quasars were about 1000 times more common than they are now. Nevertheless, even during the age of quasars, quasar eruptions must have been unusual. At any one time, only a small fraction of galaxies had quasars erupting in their cores. As expansion carried the galaxies apart, and as galaxy formation became less active, quasars became less common.

Then where are all the dead quasars? An astronomer commented recently, "There is no way to get rid of supermassive black holes, so all of the galaxies that had short-lived quasars must still have those supermassive black holes." Where are all those dead quasars today? You know where to look—the cores of galaxies.

Most galaxies contain supermassive black holes at their centers, but the black holes have gobbled up most of the nearby gas, dust, and stars, and they are dormant. Our own Milky Way Galaxy could have had a quasar at its core long ago, but now its central black hole is on a strict diet. Occasional interactions between galaxies can throw matter inward and nudge a sleeping supermassive black hole into eruption to become a Seyfert galaxy or a double-lobed radio galaxy.

Active galaxies are not some rare kind of galaxy. They are normal galaxies passing through a stage that many galaxies experience. They are not really peculiar. They are an important part of the story of the formation and evolution of galaxies.

SCIENTIFIC ARGUMENT

What do the distorted shapes of active galaxies tell you?
Seyfert galaxies and the galaxies located between double radio lobes are often distorted and have tidal tails. Also, the galaxies that are faintly visible around quasars are also typically twisted out of shape. These distortions must be caused by tidal interactions with nearby galaxies or by mergers, and such tides can throw matter inward toward the supermassive black hole at the center of a galaxy and trigger an eruption.

Now build a new scientific argument. **Why were quasars rare when the universe was young, and why are they rare today?**

What Are We? Changed

Next time you are in a shopping mall, glance at the people around you. How many of them, do you suppose, know that galaxies can collide and erupt or that there was an age of quasars? The vast majority of people have no idea how their lives fit into the story of the universe. Most people don't know what they are. They eat pizza and watch TV without understanding that they are part of a universe in which galaxy collisions trigger supermassive black holes to erupt in titanic explosions.

Astronomy is changing you. As you learn more about stars and galaxies and quasars, you are learning more about yourself and your connection with nature. "Perspective" can mean a view of things in their true relationships. As you study astronomy, you are gaining perspective. Our galaxy, our sun, our planet, and the local shopping mall take on new meaning when you think astronomically.

Study and Review

Summary

- Astronomers divide galaxies into three classes—**elliptical (p. 398)**, **spiral (p. 398)**, and **irregular (p. 399)**—with subclasses specifying the galaxy's shape.
- Elliptical galaxies contain little gas and dust and cannot make new stars, but spiral galaxies contain more gas and dust and make new stars especially in the spiral arms.
- About two-thirds of spirals, including our Milky Way Galaxy, are **barred spiral galaxies (p. 398).** Irregular galaxies have no obvious shape but contain gas and dust and young stars.
- Galaxies are so distant astronomers measure their distances in **megaparsecs (Mpc) (p. 400)**—millions of parsecs.
- Astronomers find the distance to galaxies using **distance indicators (p. 400),** sometimes called **standard candles (p. 400),** objects of known luminosity such as Cepheid variable stars, the most accurate distance indicators. By calibrating distance indicators, astronomers have built a **distance scale (p. 401).**
- When astronomers look at a distant galaxy, they see it as it was when it emitted the light now reaching Earth. The **look-back time (p. 402)** to distant galaxies can be a significant fraction of the age of the universe.
- According to the **Hubble law (p. 402),** the apparent velocity of recession of a galaxy equals its distance times the **Hubble constant (p. 402).** Astronomers can estimate the distance to a galaxy by observing its redshift, calculating its apparent velocity of recession, and then dividing by the Hubble constant.
- Once the distance to a galaxy is known, its diameter can be found from the small-angle formula and its luminosity from the magnitude-distance formula.
- The **rotation curve (p. 403)** of a galaxy shows the orbital motion of its stars, and astronomers can use the **rotation curve method (p. 403)** to find the galaxy's mass.
- Some dwarf galaxies are only a few percent the size and luminosity of our galaxy, but some giant elliptical galaxies are five times larger than the Milky Way Galaxy.
- High orbital velocities near the centers of galaxies suggest the presence of supermassive black holes.
- Rotation curve observations, hot gas held inside some clusters of galaxies, and **gravitational lensing (p. 405)** reveal the presence of dark matter.
- **Rich clusters (p. 406)** of galaxies contain thousands of galaxies with fewer spirals and more ellipticals. **Poor clusters (p. 406)** of galaxies contain few galaxies with a larger proportion of spirals. This is evidence that galaxies evolve by collisions and mergers. Our galaxy is located in a poor cluster known as the **Local Group (p. 406).**
- When galaxies collide, tides twist and distort their shapes and can produce **tidal tails (p. 410).** Large galaxies can absorb smaller galaxies in what is called **galactic cannibalism (p. 410).** Shells of stars, counterrotating parts of galaxies, streams of stars in the halos of galaxies, and multiple nuclei are evidence that galaxies can merge.
- **Ring galaxies (p. 411)** are produced by high-speed collisions in which a small galaxy plunges through a larger galaxy perpendicular to its disk.
- The compression of gas clouds can trigger rapid star formation, producing **starburst galaxies (p. 408).**
- The merger of two larger galaxies can scramble star orbits and drive bursts of star formation to use up gas and dust. Most large ellipticals have evidently been produced by past mergers. Spiral galaxies have thin, delicate disks and appear not to have suffered mergers with large galaxies.
- A galaxy moving through the gas in a cluster of galaxies can be stripped of its own gas and dust and may become an S0 galaxy.
- First called **radio galaxies (p. 409), active galaxies (p. 409)** are now known to emit energy at many wavelengths. Because the energy comes from their cores, they are known also as **active galactic nuclei (AGN) (p. 409).**
- **Seyfert galaxies (p. 409)** are spiral galaxies with small, highly luminous cores and spectra that show the nuclei contain highly excited gas.
- **Double-lobed radio galaxies (p. 412)** emit radio energy from areas on either side of the galaxies. As described by the **double-exhaust model (p. 414),** these lobes are inflated by jets ejected from the nuclei of the galaxies. Where the jets run into surrounding gas, they form **hot spots (p. 414).**
- The **quasars (p. 414)** have a starlike appearance but very large redshifts that imply they are very distant. To be visible at such distances, quasars must be ultraluminous. Yet rapid fluctuations in brightness show that quasars must be small, with diameters of only a few light hours, comparable in size to our solar system.
- The best images show that quasars are embedded in galaxies that are often distorted or have close companions. From this, astronomers conclude that quasars are the highly luminous cores of very distant active galaxies.
- Matter flowing into a supermassive black hole must conserve angular momentum and form an accretion disk. A few such disks have been imaged in the cores of active galaxies, and measurement of their rotation velocities permit determination of the masses of the black holes.
- The spinning accretion disk pulls in and winds up the magnetic field, and by a process not yet fully understood, this magnetic field ejects and focuses two jets of radiation and high-speed gas in opposite directions along the axis of rotation.
- Supermassive black holes erupt only when large amounts of matter flow inward, so most galaxies have dormant cores with little mass falling inward. During galaxy interactions and collisions, tides can throw matter into the galaxy centers and trigger eruptions.
- Because quasars lie at great distances, astronomers see them as they were long ago—over 10 billion years ago—when the universe was young and just forming galaxies.
- Astronomers are developing a **unified model (p. 418)** that explains differences among some AGN as differences in the inclination of the accretion disk.
- Because the mass of a supermassive black hole is related to the mass of the galaxy central bulge around it, astronomers conclude that the black holes formed when the first gas clouds fell together to form the central bulges.
- Mergers with smaller galaxies are thought to have built the halos of galaxies and may have triggered eruptions of the cores. However, the formation of the large disks and spiral arms that lie outside the central bulges occurred later as gas settled into the galaxies.
- Some of the most distant quasars may be erupting because of the formation of supermassive black holes, but most activity is triggered by interactions and collisions with other galaxies. Collisions were more common in the past before the universe had expanded very far and separated galaxies.
- The age of quasars, at redshifts from about 2 to 3, was a time when galaxies were actively growing and merging.
- The supermassive black holes that once produced quasar eruptions are now mostly dormant because very little mass is flowing into them. A galaxy can be triggered to become an active galaxy if new mass falls into the black hole.

Review Questions

1. Why can't the evolution of galaxies go from elliptical to spiral? From spiral to elliptical?
2. What is the difference between an Sa and an Sb galaxy? Between an S0 and an Sa galaxy? Between an Sb and an SBb galaxy? Between an E7 and an S0 galaxy?
3. Explain how the rotation curve method of finding a galaxy's mass is similar to the method used to find the masses of binary stars.
4. Explain how the Hubble law allows you to estimate the distances to galaxies.
5. How can collisions affect the shapes of galaxies?
6. What evidence can you cite that galactic cannibalism really happens?
7. What evidence suggests that Seyfert galaxies have suffered recent interactions with other galaxies?
8. How does the peculiar rotation of NGC 5128 help explain the origin of this active galaxy?
9. What evidence shows that quasars are ultraluminous but must be small?
10. What evidence is there that quasars occur in distant galaxies?
11. Why are there few quasars at low redshifts and at very high redshifts but many at redshifts between about 2 and 3?
12. **How Do We Know?** Classification helped Darwin understand how creatures evolve. Has classification helped you understand how galaxies evolve?
13. **How Do We Know?** How would you avoid selection effects if you were taking a survey of the kinds of wild flowers found in a meadow?
14. **How Do We Know?** How would you respond to someone who said, "Oh, that's only statistics"?

Discussion Questions

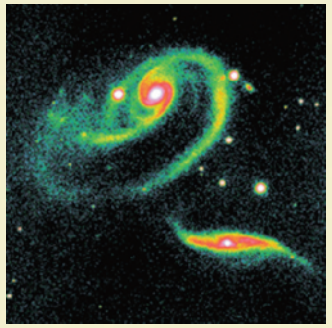
1. From what you know about star formation and the evolution of galaxies, do you think irregular galaxies should be bright or faint in the infrared relative to visible wavelengths? Why or why not? What about starburst galaxies? What about elliptical galaxies?
2. Imagine that you could observe a few gas clouds at such a high look-back time that they are just beginning to form one of the first galaxies. Further, suppose you discovered that the gas was metal rich. Would that support or contradict your understanding of galaxy formation?
3. Do you think that our galaxy is currently an active galaxy? Do you think it ever was in the past? Could it have hosted a quasar when it was young?

Problems

1. If a galaxy contains a type I (classical) Cepheid with a period of 30 days and an apparent magnitude of 20, what is the distance to the galaxy? (*Hints:* See Figure 17-4 and Reasoning with Numbers 13-2.)
2. If a galaxy contains a supernova that at its brightest has an apparent magnitude of 17, how far away is the galaxy? Assume that the absolute magnitude of the supernova is −19. (*Hint:* See Reasoning with Numbers 13-2.)
3. If you find a galaxy that is the same size and mass as our Milky Way Galaxy, what orbital velocity would a small satellite galaxy have if it orbits 50 kpc from the center of the larger galaxy?
4. If a galaxy has an apparent radial velocity of 2000 km/s and the Hubble constant is 70 km/s/Mpc, how far away is the galaxy?
5. Suppose you found a galaxy in which the outer stars have orbital velocities of 150 km/s. If the radius of the galaxy is 4.0 kpc, what is the orbital period of the outer stars? (*Note:* 1 pc = 3.1×10^{13} km, and 1 yr = 3.2×10^{7} seconds.)
6. Among the globular clusters orbiting a distant galaxy, one is moving at 420 km/s and is located 11 kpc from the center of the galaxy. Assuming the globular cluster is located outside most of the mass of the galaxy, what is the mass of the galaxy? (*Hints:* Use the formula for circular velocity, Chapter 4. Remember to convert units to meters and kilograms.)
7. If the jet in NGC 5128 is traveling at 5000 km/s and is 40 kpc long, how long will it take for gas to travel from the core of the galaxy to the end of the jet? (*Note:* 1 pc equals 3×10^{13} km.)
8. If the active core of a galaxy contains a black hole of 10^{6} solar masses, what will the orbital period be for matter orbiting the black hole at a distance of 0.33 AU? (*Hint:* See Reasoning with Numbers 4-1.)
9. If a quasar is 1000 times more luminous than an entire galaxy, what is the absolute magnitude of such a quasar? Note, the absolute magnitude of a bright galaxy is about −21. (*Hint:* See Reasoning with Numbers 2-1.)
10. The hydrogen Balmer line H_β has a wavelength of 486.1 nm. It is shifted to 563.9 nm in the spectrum of 3C 273. What is the redshift of this quasar? (*Hint:* What is $\Delta\lambda$?)

Learning to Look

1. Hubble's first determination of the Hubble constant was too high because the calibration of Cepheid variable stars he used was not very good. Use his original diagram shown in Figure 18-5 to estimate his first determination of the Hubble constant.
2. In the image at right you see two interacting galaxies; one is nearly face-on and the other is nearly edge-on. Discuss the shapes of these galaxies and describe what is happening.

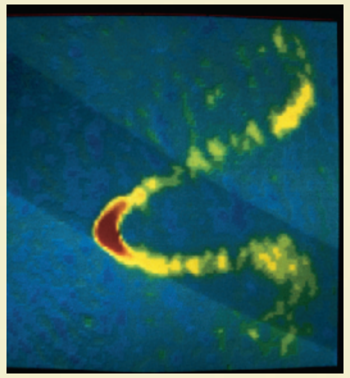


WIN/NOAO/NSF

3. The image at right combines visual (blue) with radio (red) to show the galaxy radio astronomers call Fornax A. Explain the features of this image. Is it significant that the object is a distorted elliptical galaxy in a cluster?

NRAO/AUI

4. Explain the features of this radio image of the galaxy IC 708.



NRAO/AUI

Great Debates

1. *Are You An Alien?* An alien to Earth is life not of Earth. Some hypotheses suggest that ice was delivered to Earth by comet storms from the Oort cloud during previous collisions of the Milky Way Galaxy with other galaxies in our Local Group. The galaxy collisions scrambled star orbits and passing stars could distort each other's Oort Clouds, mixing comet populations and throwing comets inward toward the planets. If this is true then some of the water molecules in your body may have originated in another galaxy. Does this mean you are an alien to Earth?
   a. Use at least three vocabulary words from the text correctly, underline each in your debate, and cite page and paragraph numbers in your answer.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
2. *Recycled Humans?* Each time a star lives and dies, the star system returns matter enriched in metals to the interstellar medium. If humans are recycled matter, is matter content today different from our matter content in the past? If humans are recycled, will our matter content be different in the future as compared to today? Will humans someday be silicon based instead of water based?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
3. *Hubble's Law.* In 1929, Hubble published a paper describing the relationship between distances to galaxies derived from Cepheids and the speeds at which these galaxies are receding from us as determined from the redshifts in their spectra. This relationship was also published two years prior in an obscure Belgian journal. The author of the 1927 paper is Georges Lemaitre, a Catholic priest and physicist. In 1931, Lemaitre was invited to translate his earlier paper into English and publish it in the world's leading journal at that time. Lemaitre chose to not translate key portions of his original paper because newer evidence was already in Hubble's paper. Nonetheless, the original idea of observations supporting the expansion of the universe is in the 1927 paper by Lemaitre. Should the law be renamed to the Hubble–Lemaitre Law or the Lemaitre–Hubble Law?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
4. *Galaxy Definition.* Recently the definition of galaxy has come under fire. As demonstrated in this text, galaxies come in all kinds of shapes, sizes, and colors, and dark matter is invoked to explain some observations. The word *galaxy* has no standard definition. Should the International Astronomical Union call for a vote on the definition of *galaxy* like the vote on the definition of *planet* that ultimately led to Pluto's reclassification as a dwarf planet?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
5. *Feline Galaxy Guardians.* In *Men in Black I*, Agent J searches for the guardian of a galaxy (that is, the cat) to find and secure the galaxy before alien bugs find it and destroy Earth in the process. Is a galaxy the size of a pendant on a cat's collar realistic?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Variable Stars in a Galaxy
- Gravitational Lensing and Dark Matter
- Formation of Galaxy Shells
- Active Galaxy Jets
- Hubble Diagram

CengageNOW Virtual Astronomy Labs 2.0

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 19: The Hubble Law

You obey the law against littering out of civic pride but also to avoid getting arrested. Is it right to say the universe "obeys" the Hubble law? No one will give a galaxy a ticket if it breaks that law. In science, a natural law is not really obeyed; rather, that word is used to describe how nature works. Scientists use the word *law* to indicate those descriptions of nature that are the most reliable, true at all times and in all places.

As you learned in this chapter, the Hubble law describes a simple relationship between the distances of galaxies and their redshifts. The law was originally discovered by astronomer Edwin Hubble in the 1920s and named after him by his colleagues. Hubble's primary method of determining distances was to identify Cepheid variable stars in other galaxies, then to measure their brightness and pulsation periods. Cepheids in distant galaxies could then be compared with nearby Cepheids using the inverse-square law, as you

studied in the previous chapter. Finally, Hubble needed to take spectra of those galaxies to measure the redshifts of their spectral lines. Today, astronomers make the same measurements of galaxies far more distant than Hubble the person could reach using *Hubble* the telescope. In fact, the *Hubble Space Telescope* was named after Edwin Hubble because its first goal after being launched into orbit around Earth was to carry on his work and make a refined determination of the Hubble law.

Looking back from your present extragalactic vantage point, you can see clearly that determination of distances to faraway galaxies depends on a series of techniques and measurements that lead upward and outward like steps on a stairway or rungs on a ladder. In fact, astronomers often call this the "Distance Ladder." Going back down from the top of the ladder to the bottom: (4) Extragalactic distances depend on inverse square law analyses of observations of Cepheid variable stars (and other "standard candles"), which are (3) calibrated by measurements of Cepheids near enough that their distances are known from parallax and other geometric methods, which in turn (2) use the radius of Earth's orbit—1 AU—as the trigonometric baseline, and finally, at the bottom, (1) the size of Earth's orbit is known from knowledge of Kepler's laws plus parallax measurements of the distances to the other planets using the size of Earth as a baseline. The known size of Earth, which we have sailed and flown and ridden and walked around thoroughly, is the ultimate foundation of the extragalactic distance scale.

As you learned in this chapter, quasars are understood to be the most luminous examples of AGN (active galactic nuclei). Some have very large redshifts and are among the most distant objects known. Don't be misled, however, by the common practice of astronomers expressing redshifts of distant galaxies and quasars as radial velocities in kilometers per second. Those are only apparent radial velocities, not real velocities through space. This important point will be explored further in the following chapter.

Section 1 of Virtual Astronomy Lab 19, "The Hubble Law," helps you make your own determination of the Hubble constant (also known as the Hubble parameter)—the proportionality between distances and redshifts. In Section 2 you can make measurements of a quasar's spectrum to determine its redshift and distance. Knowing the distance of quasars unlocks the mysteries of those powerful objects. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 18: Active Galactic Nuclei

Galaxies are incomprehensibly huge; diameters of 100,000 ly are not unusual. It is startling, then, to realize that something so large can erupt violently, producing powerful jets. As you learned in this chapter, astronomers have multiple lines of evidence that trace the ultimate cause of those eruptions back to the supermassive black holes lying in the centers of most large galaxies. You can investigate these phenomena for yourself. First, gather information. What are the different kinds of active galaxies? What are the eruptions and jets like? Those are the basic data you need to understand what happens when an active galaxy erupts.

Scientists carefully distinguish between what is observed and what is really happening. Some jets rushing away from active galaxies appear to exceed the speed of light. Einstein's theory of special relativity directly implies that nothing can go faster than light. Special relativity has been checked many times and in many ways and always found to be exactly correct. So what is really happening in those so-called superluminal (meaning "above light-speed") jets? Are those jets providing clues that might let humanity invent "warp drive" and zoom through the universe faster than light?

In Sections 1 and 2 Virtual Astronomy Lab 18, "Active Galaxy Nuclei," an animation and an exercise will help you understand that, unfortunately for the warp drive scenario, the observations of apparently faster-than-light motions are just illusions produced by the high speeds of the jets and the angle at which they are approaching Earth. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 17: Evidence for Dark Matter

In the previous chapter you learned about how evidence for existence of dark is apparent if you make a plot of the orbital velocity of stars and gas clouds versus their distance from the center of our Milky Way Galaxy. Such a plot is called a rotation curve. The galaxy's rotation curve is approximately constant ("flat") for distances beyond the radius of the sun's orbit, showing the gravitational effect of dark matter in the Milky Way. In this chapter you learned that the same phenomenon is observed for many other large galaxies. Their rotation curves also indicate large amounts of matter extending to great distances from each galaxy's center that exerts gravitational force but does not emit or absorb any form of electromagnetic radiation. In this chapter you also saw evidence that entire clusters of galaxies show signs of widespread dark matter, based on a calculation of the amount of gravity needed to retain very hot gas detected by X-ray telescopes within the clusters.

If you want further evidence of dark matter, you can consider the phenomenon called gravitational lensing, which is described in this chapter. When light from a very distant galaxy passes close to a nearer cluster of galaxies, the mass of the foreground galaxy cluster curves space-time and deflects the light on its way to Earth. That means you observe the light from the distant galaxy to be arriving here from two directions, making the distant background galaxy appear as galaxies located on opposite sides of the foreground galaxy cluster. Those images are only mirages, but measurement of their locations can tell you the total mass in the foreground galaxy cluster necessary to cause the lensing effect. If you subtract the mass of the visible galaxies, you can show that the galaxy cluster must contain a large amount of dark matter. You should note that this method of detecting dark matter is completely independent of the other methods described previously. The different methods agree about the amount of dark matter, confirming the overall picture.

Section 2 of Virtual Astronomy Lab 17, "Evidence for Dark Matter," contains an exercise that helps you make the connection between observations of gravitational lensing and the mass of dark matter contained in clusters of galaxies. There is a redshift calculator supplied with the exercise that adds interesting information but is not necessary to perform the lensing exercise itself. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Tutorial 13.4 – Cosmic Rays and Active Galactic Nuclei

Almost all the information we have about celestial objects comes from analysis of light and other varieties of electromagnetic radiation received on Earth from them. Almost all. Humanity has "in our hands", so to speak, physical pieces of the moon, Mars, asteroids Vesta and Itokawa, meteorites and micrometeorites that are fragments of asteroids with names unknown, dust gathered from Comet Wild 2's tail, and samples of gas atoms from the solar wind. Also, *in situ* measurements have been made by robot probes of the properties of Luna's surface, Venus's, Mars's, and Titan's surfaces and atmospheres, Jupiter's atmosphere, and the magnetospheres of most of the planets and many of their moons. Neutrinos produced in the core of the sun have been detected and counted.

What about physical evidence from farther away than our solar system? The collector on the *Stardust* spacecraft appears to have obtained samples of interstellar dust in addition to comet dust, and grains suspected of being interstellar, pre-dating the solar system, have been found embedded in some meteorites. Bursts of neutrinos from Supernova 1987A were detected by several facilities before the light of the explosion had reached Earth.

You might think that supernova neutrinos traveling 50,000 parsecs from the Large Magellanic Cloud in the distant halo of the Milky Way Galaxy would be the record holders. But, some cosmic rays, atomic nuclei traveling nearly at the speed of light, have been detected with such high energies that their paths would not have been deflected even by passage through the entire Milky Way Galaxy's magnetic field. Thus, the directions from which they are coming may indicate their true origins in Active Galactic Nuclei (AGN) millions of light years distant.

Section 4 of Virtual Astronomy Lab 4, "Solar Wind and Cosmic Rays," guides you through understanding how cosmic rays are detected and how some may be produced by supernova explosions, while others may be coming to us directly from the unimaginably violent centers of AGN. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

19

Modern Cosmology

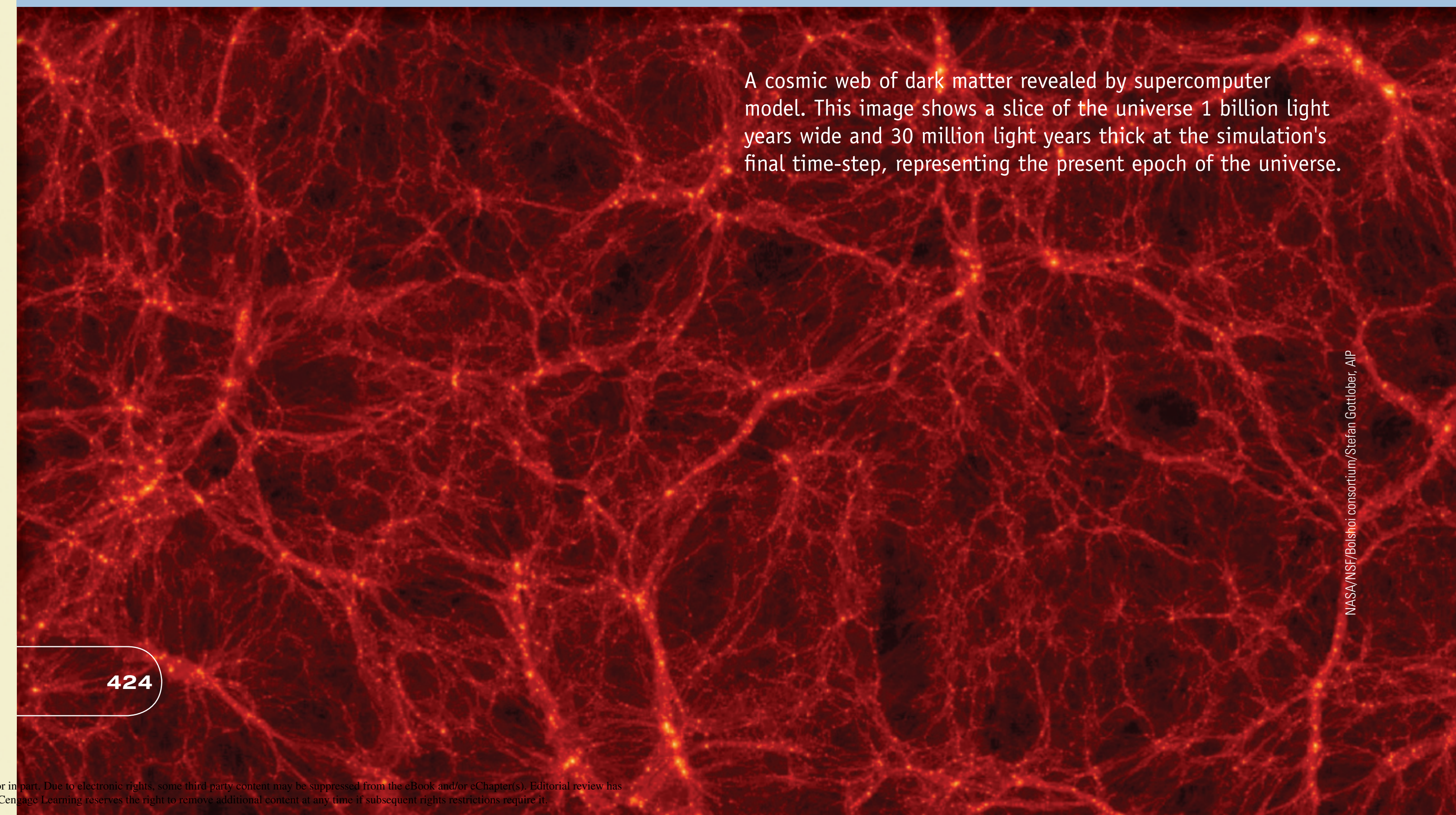
Guidepost

Since Chapter 1, you have been on an outward journey through the universe. Now you have reached the limit of your travels in space and time and can contemplate the universe as a whole. The ideas in this chapter are among the biggest and most difficult in all of science. Can you imagine a limitless universe, or the first instant of time?

As you explore cosmology, you will find answers to three important questions:

- **Does the universe have a center and an edge?**
- **What is the evidence that the universe began with a big bang?**
- **How has the universe evolved, and what will be its fate?**

Once you have finished this chapter, you will have a modern insight into the nature of the universe, as well as where you are and what you are.



A cosmic web of dark matter revealed by supercomputer model. This image shows a slice of the universe 1 billion light years wide and 30 million light years thick at the simulation's final time-step, representing the present epoch of the universe.

NASA/NSF/Bolshoi consortium/Stefan Gottlöber, AIP

If the doors of perception were cleansed everything would appear to man as it is, infinite.

WILLIAM BLAKE

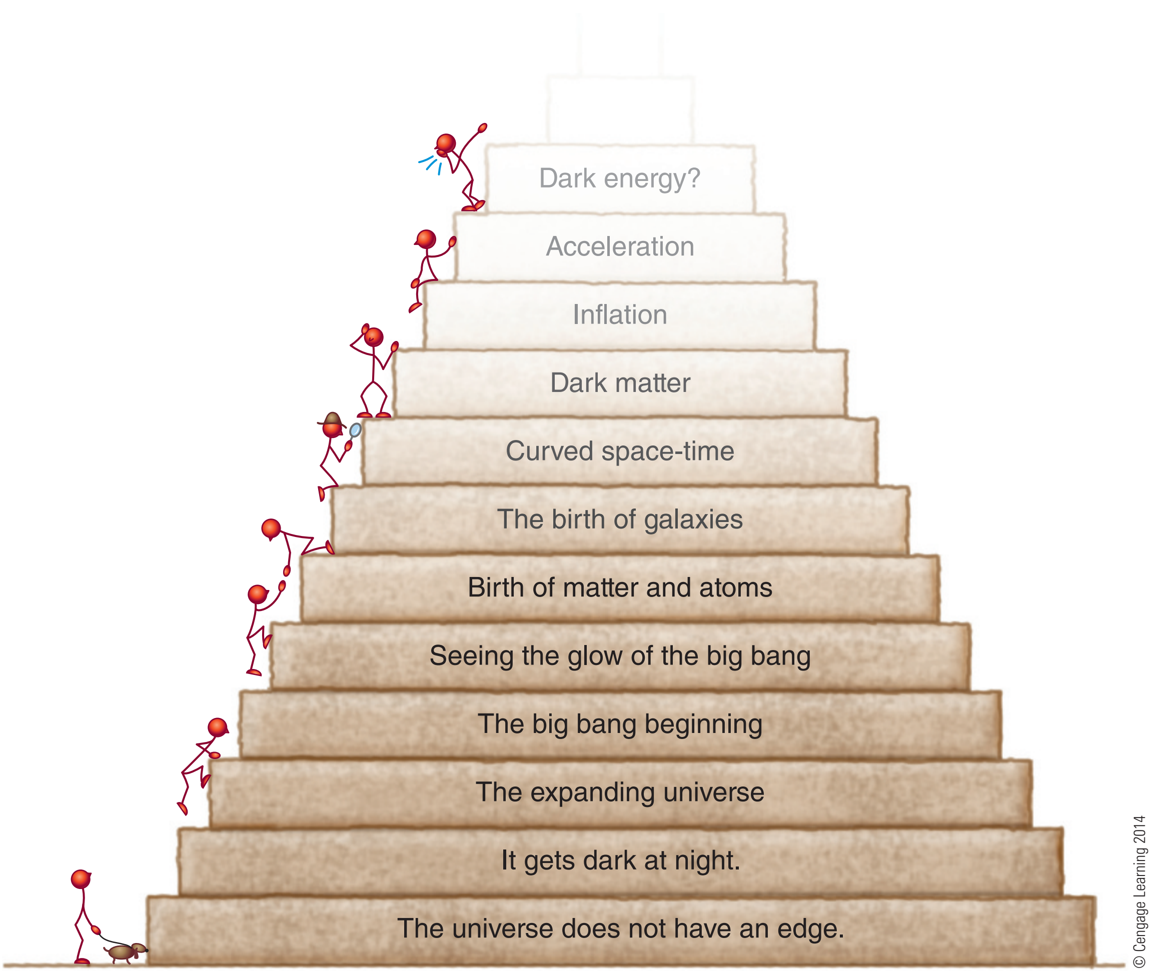
LOOK AT YOUR THUMB. The matter in your body was present in the fiery beginning of the universe. **Cosmology,** the study of the universe as a whole, can tell you where your matter came from, and it can tell you where your matter is going.

Cosmology is a mind-bendingly weird subject, and you can enjoy it for its strange ideas. It is fun to think about space stretching like a rubber sheet, invisible energy pushing the universe to expand faster and faster, and the origin of vast walls of galaxy clusters. Notice that this is better than speculation—it is all supported by evidence. Cosmology, however strange it may seem, is a serious and logical attempt to understand the structure and evolution of the entire universe.

This chapter will help you climb the cosmology pyramid one step at a time (■ Figure 19-1). You already have some ideas about what the universe is like. Start with those ideas, test them against observations, and also compare them with scientific theories. Step-by-step you can build a modern understanding of cosmology. Each step in the pyramid is small, but it leads to some astonishing insights into how the universe works and how you came to be a part of it.

19-1 Introduction to the Universe

MOST PEOPLE HAVE AN IMPRESSION of the universe as a vast ocean of space filled with stars and galaxies (■ Figure 19-2), but as you begin exploring the universe, you need to become aware of your expectations so they do not mislead you. The first step is to deal



■ **Figure 19-1**

Climbing the cosmology pyramid step by step isn't very difficult, and it leads to some fascinating ideas about the origin and evolution of the universe.

R. Williams, STScI HDF-South Team, NASA, ESA

Figure 19-2

The entire sky is filled with galaxies. Some lie in clusters of thousands, and others are isolated in nearly empty voids between the clusters. In this image of a typical spot on the sky, the brightest objects are nearby stars; their "spikes" are caused by diffraction in the telescope. All other objects are galaxies ranging from the relatively close face-on spiral at upper right to the most distant galaxies visible only in the infrared, shown as red in this composite image.

with an expectation so obvious that most people don't think about it, for the sake of a quiet life.

The Edge–Center Problem

In your daily life you are accustomed to boundaries. Rooms have walls, athletic fields have boundary lines, countries have borders, oceans have shores. It is natural to think of the universe also as having an edge, but that idea can't be right.

If the universe had an edge, imagine going to that edge. What would you find there: A wall of some type? A great empty space? Nothing? Even a child can ask: If there is an edge to the universe, what's beyond it? A true edge would have to be more than just an end of the distribution of matter. It would have to be an end of space itself. But, then, what would happen if you tried to reach past, or move past, that edge?

An edge to the universe violates common sense, and modern observations (which you will study later in this chapter) indicate that the universe could be infinite and would therefore have no edge. Perhaps even more important, if the universe has no edge, then it cannot have a center. You find the centers of things—pizzas, football fields, oceans, galaxies—by referring to their edges. If the universe has no edge, then it cannot have a center.

It is a **Common Misconception** to imagine that the universe has a center, but, as you just realized, that is impossible. As you study cosmology, you should take care to avoid thinking that there is a center of the universe.

The Idea of a Beginning

Of course you have noticed that the night sky is dark. That is an important observation because, you may be surprised to learn, seemingly reasonable assumptions about the universe can lead to the conclusion that the night sky actually should glow blindingly bright. This conflict between observation and theory is called **Olbers's paradox** after Heinrich Olbers, a physician and astronomer who publicized the problem in 1826. (*Problem* or *question* might be more accurate words than *paradox.*) The problem of the dark night sky was first discussed by Thomas Digges in 1576, but Olbers gets the credit through an accident of scholarship on the part of modern scientists who were not aware of earlier discussions.

The point made by Olbers seems simple. Suppose you assume, as did most scientists in Olbers's time, that the universe is infinite in size, infinite in age, **static** (a fancy word for unchanging overall), and filled with stars. If you look in any direction, your line of sight must eventually run into the surface of a star. (The clumping of stars into galaxies and galaxies into clusters can be shown mathematically to make no difference.) Look at ■ Figure 19-3, which uses the analogy of lines of sight in a forest. (The use of analogies in science is discussed in **How Do We Know? 19-1**). When you are deep in a forest, every line of sight ends at a tree trunk, and you cannot see out of the forest.

By analogy to the view from inside a forest, every line of sight from Earth into space should eventually end at the surface of a star. Of course, the more distant stars would be fainter than nearby stars because of the inverse square law. However, the farther you look into space, the larger the volume you are viewing and the more stars are included; the two effects cancel out. The result would be that the entire sky should be as bright as the surface of an average star—like suns crowded "shoulder to shoulder," covering the sky from horizon to horizon. It should not get dark at night.

Astronomers and physicists who study cosmology, called **cosmologists,** now believe they understand why the sky is dark. Olbers's paradox makes an incorrect prediction because it is based on incorrect assumptions. The universe may be infinite in size, but it is neither infinitely old nor static. The essence of modern cosmologists' answer to Olbers's question was suggested first by Edgar Allan Poe in 1848. Poe proposed that the night sky is dark because the universe is not infinitely old but came into existence at some finite time in the past. The more distant stars are so far away that light from them has not yet reached Earth. That is, if you look far enough away, the look-back time approaches the age of the universe, and you see to a time before the first stars

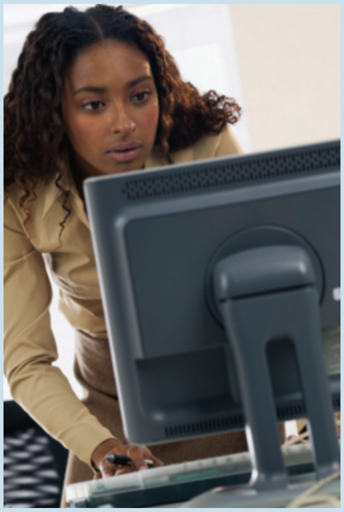
Reasoning by Analogy

How do scientists use analogies? "The economy is overheating, and it may seize up," an economist might say. Economists like to talk in analogies because economics is often abstract, and one of the best ways to think about abstract problems is to find a more approachable analogy. Rather than discussing details of the national economy, you might be able to make conclusions about how the economy works by thinking about how a gasoline engine works.

Much of astronomy is abstract, and cosmology is the most abstract subject in astronomy. Furthermore, cosmology is highly mathematical, and unless you are prepared to learn some difficult mathematics, you instead have to use analogies, such as lines of sight in a forest.

Reasoning by analogy is a powerful technique. An analogy can reveal unexpected insights and lead you to further discoveries. Carrying an analogy too far, however, can be misleading. You might compare the human brain to a computer, and that would help you understand how data flow in and are processed and how new data flow out, but the analogy is flawed. For example, although data in computers are stored in specific locations, memories are stored in the brain in a distributed form. No single brain cell holds a specific memory. So, if you carry the analogy too far, it can mislead you. Whenever you reason using analogies, you should be alert for their limitations.

As you study any science, be alert for analogies. They are tremendously helpful, but you have to be careful not to carry them too far.



Blend Images/SuperStock

The analogy between a human brain and a computer is of only limited use.

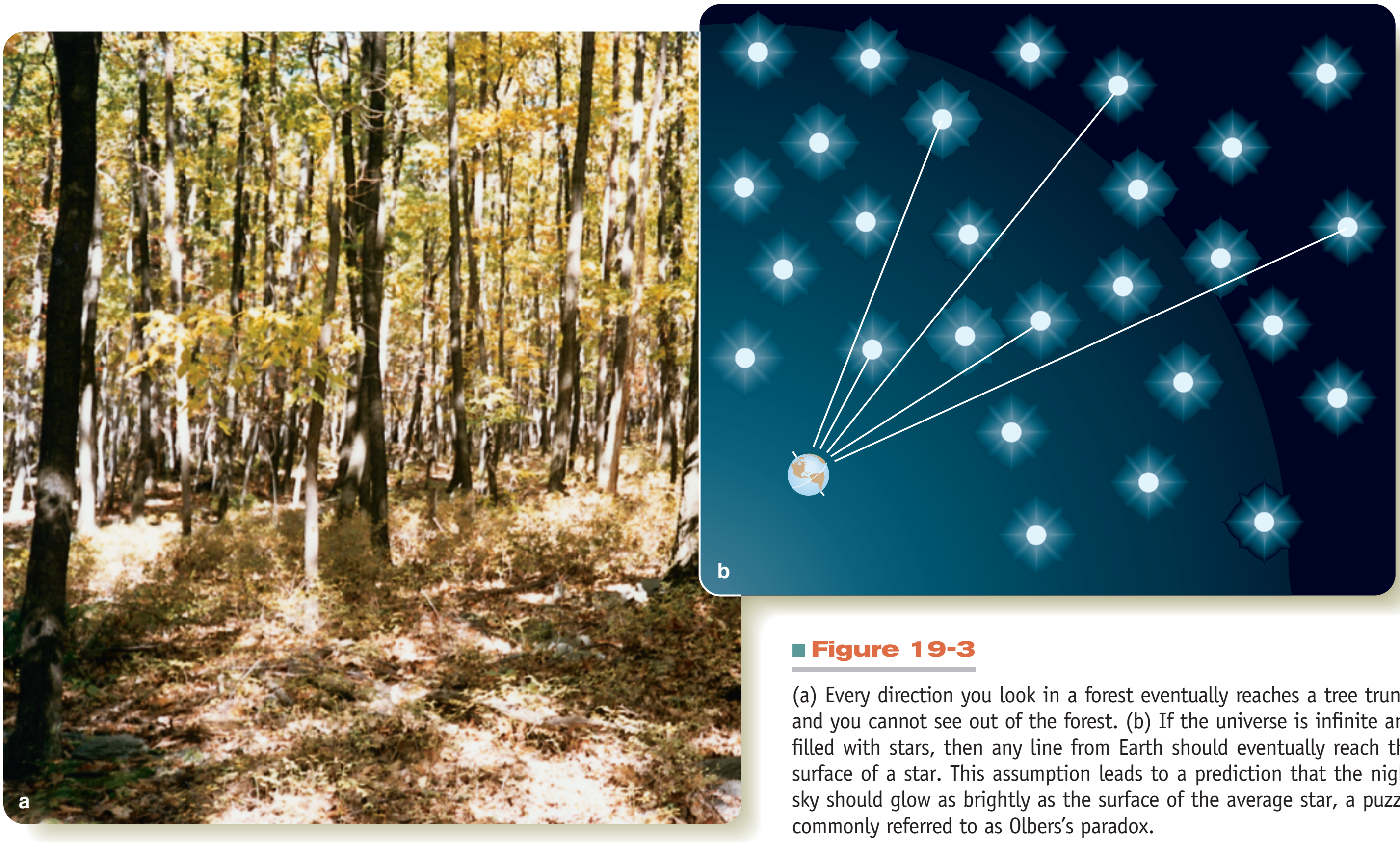


Photo courtesy of Janet Seeds

© Cengage Learning 2014

Figure 19-3

(a) Every direction you look in a forest eventually reaches a tree trunk, and you cannot see out of the forest. (b) If the universe is infinite and filled with stars, then any line from Earth should eventually reach the surface of a star. This assumption leads to a prediction that the night sky should glow as brightly as the surface of the average star, a puzzle commonly referred to as Olbers's paradox.

began to shine. The night sky is dark because the universe had a beginning. You can now answer Olbers's question, and understand why the night sky is dark, by revising your original assumptions about the universe.

The answer to Olbers's question is a powerful idea because it clearly illustrates the difference between the universe and the **observable universe.** The universe is everything that exists, and it could be infinite. The observable universe, in contrast, is the part (maybe a very small part) that you can see from Earth using the most powerful telescopes. You will learn later in this chapter about evidence that the universe is around 14 billion years old. That means you can't observe objects farther away than a look-back time of around 14 billion years. Do not confuse the observable universe, which is huge but finite, with the universe as a whole, which might be infinite.

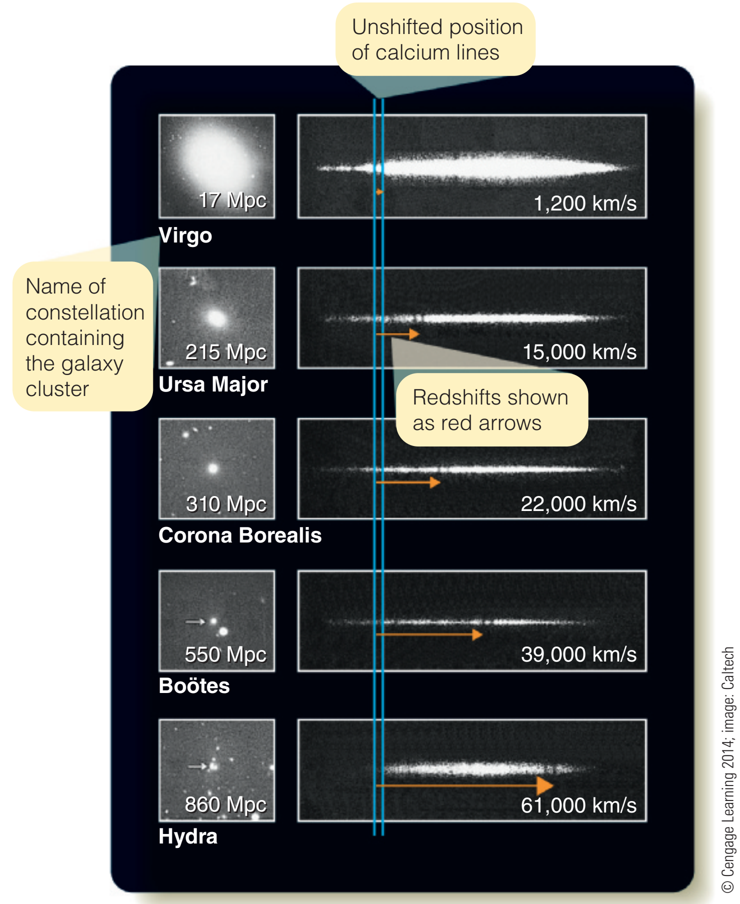
There is a **Common Misconception** that the universe is unchanging overall and not evolving. In the next section you will discover that the universe is actually changing.

Cosmic Expansion

In 1929, Edwin Hubble published his discovery that the sizes of galaxy redshifts are proportional to galaxy distances. Nearby galaxies have small redshifts, and more distant galaxies have large redshifts. You learned this as the Hubble law in Chapter 18, when you used it to estimate the distances to galaxies. Those galaxy redshifts, interpreted as Doppler shifts, imply that galaxies are receding from each other, an idea that became known as the **expanding universe.**

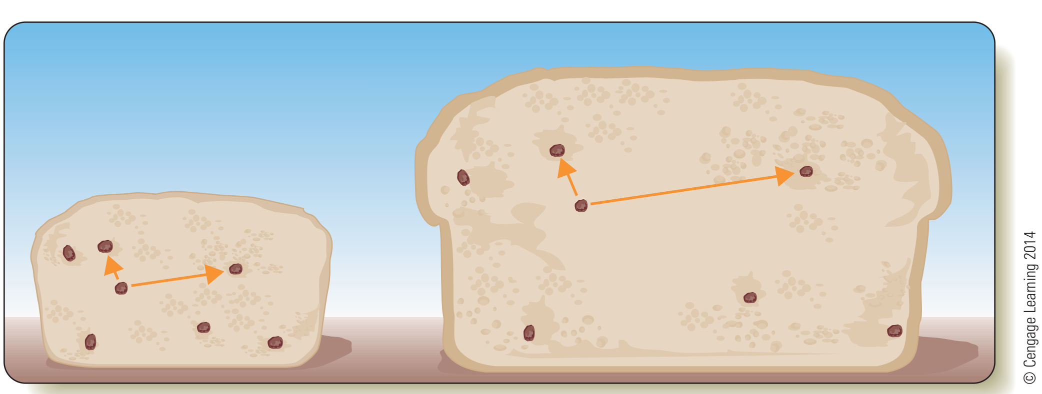
■ Figure 19-4 shows spectra of galaxies in galaxy clusters at various distances. The Virgo cluster is relatively nearby, and its redshift is small. The Hydra cluster is very distant, and its redshift is so large that the two dark spectral lines formed by ionized calcium are shifted from near-ultraviolet wavelengths well into the visible part of the spectrum.

The expansion of the universe does not imply that Earth has a special location. To see why, look at ■ Figure 19-5, which shows an analogy to baking raisin bread. As the dough rises, it pushes the raisins away from each other at speeds that are proportional to distance. Two raisins that were originally close to each other are pushed apart slowly, but two raisins that were far apart, having more dough between them, are pushed apart faster. If bacterial astronomers lived on a raisin in the raisin bread, they could observe the redshifts of the other raisins and derive a bacterial Hubble law. They would conclude that their universe was



■ **Figure 19-4**

These galaxy spectra extend from the near-ultraviolet at left to the blue part of the visible spectrum at right. The two dark absorption lines of once-ionized calcium are prominent in the near-ultraviolet. The redshifts in galaxy spectra are expressed here as apparent velocities of recession. Note that the apparent velocity of recession is proportional to distance, which is known as the Hubble law.



■ **Figure 19-5**

An illustration of the raisin bread analogy for the expansion of the universe. As the dough rises, raisins are pushed apart with velocities proportional to distance. A colony of bacteria living on any raisin will find that the redshifts of the other raisins are proportional to their distances.

expanding uniformly. It does not matter which raisin the bacterial astronomers lived on, they would get the same Hubble law—no raisin has a special viewpoint. Similarly, astronomers in any galaxy will see the same law of expansion—no galaxy has a special viewpoint.

When you look at Figure 19-5, you see the edge of the loaf of raisin bread, and you can identify a center to the loaf. The raisin bread analogy to the universe stops working when you consider the crust (the edge) of the bread. Remember that the universe cannot have an edge or a center, so there is no center to the expansion. The raisin bread analogy is useful but also imperfect.

19-2 The Big Bang Theory

Now you are ready to take a historic step up the cosmology pyramid. The expansion of the universe led cosmologists to conclude that the universe must have begun with an event of cosmic intensity.

Necessity of the Big Bang

Imagine that you have a video of the expanding universe and run it backward. You would see the galaxies getting closer to each other. There is no center to the expansion of the universe, so you would not see galaxies approaching a single spot. Rather, you would see the spaces between galaxies shrinking and the distances between all galaxies decreasing. Eventually, as your video ran farther back, galaxies would begin to merge. If you ran the video far enough back, you would see the matter and energy of the universe compressed into a high-density, high-temperature state. The expanding universe must have begun from this moment of extreme conditions that cosmologists call the **big bang.**

How long ago did the universe begin? You can estimate the age of the universe with a simple calculation. If you need to drive to a city 100 miles away, and you can travel 50 miles per hour, you divide distance by rate of travel and learn the travel time—in this example, 2 hours. In a similar fashion, to find the age of the universe, you can divide the distance between two galaxies by the speed with which they are moving away from each other and find out how much time they have taken to reach their present separation. This simple procedure gives an estimate of the age of the universe known as the **Hubble time,** which is the same no matter which two galaxies you pick. The details of this calculation are given in **Reasoning with Numbers 19-1**.

You will fine-tune your estimate of the age of the universe later in this chapter, but for the moment you can conclude that basic observations of the universe, especially the recession of the galaxies, require that the expansion of the universe began about 14 billion years ago.

Reasoning with Numbers | 19-1

The Age of the Universe

Dividing the distance to a galaxy by the apparent velocity with which it recedes gives you an estimate of the age of the universe, and the Hubble constant simplifies your task further. The Hubble constant H has the units km/s per Mpc, which is a velocity divided by a distance. If you calculate $1/H$, you have a distance divided by velocity. To finish the division and get an age, you need to convert megaparsecs into kilometers, and then the distances will cancel out and leave you with an age in seconds. To get years, you divide by the number of seconds in a year. If you make these simple changes in units, the age of the universe in years is approximately 10^{12} divided by H in its normal astronomical units, km/s/Mpc:

$$\tau_H = \frac{10^{12}}{H} \text{ years}$$

This estimate of the age of the universe is known as the Hubble time. For example, if H is 70 km/s/Mpc, that leads to an estimated age for the universe of $10^{12}/70$, or 14 billion, years.

The phrase *"big bang"* was invented by early critics of that theory, and the label gives a misimpression. Do not think of an edge or a center when you think of the big bang. It is a very **Common Misconception** that the big bang was an explosion and that the galaxies are flying away from the location of that explosion. Instead, you should keep firmly in mind the correct picture that the big bang did not occur at a single place but filled the entire volume of the universe. A more accurate term than *big bang* might be *big stretch.* The matter of which you are made was part of that big bang, so you are inside the remains of that event, and the universe continues to expand around you. You cannot point to any particular place and say, "The big bang occurred over there." This is a brain-straining (stretching?) idea, but you will find more information to help you understand it better later in this chapter when you study the nature of space and time.

Your instinct also is to think of the big bang as an event that happened long ago and can no longer be observed, like the Gettysburg Address. Amazingly, the effect of look-back time makes it possible to observe the big bang now, directly. The look-back time to nearby galaxies is "only" a few million years; the look-back time to more distant galaxies is a large fraction of the age of the universe (■ Figure 19-6).

Suppose you look between the distant galaxies, seeing even farther away and farther back in time. You should be able to detect the hot gas that filled the universe long ago, right after the big bang, before the first stars and galaxies formed. The big bang

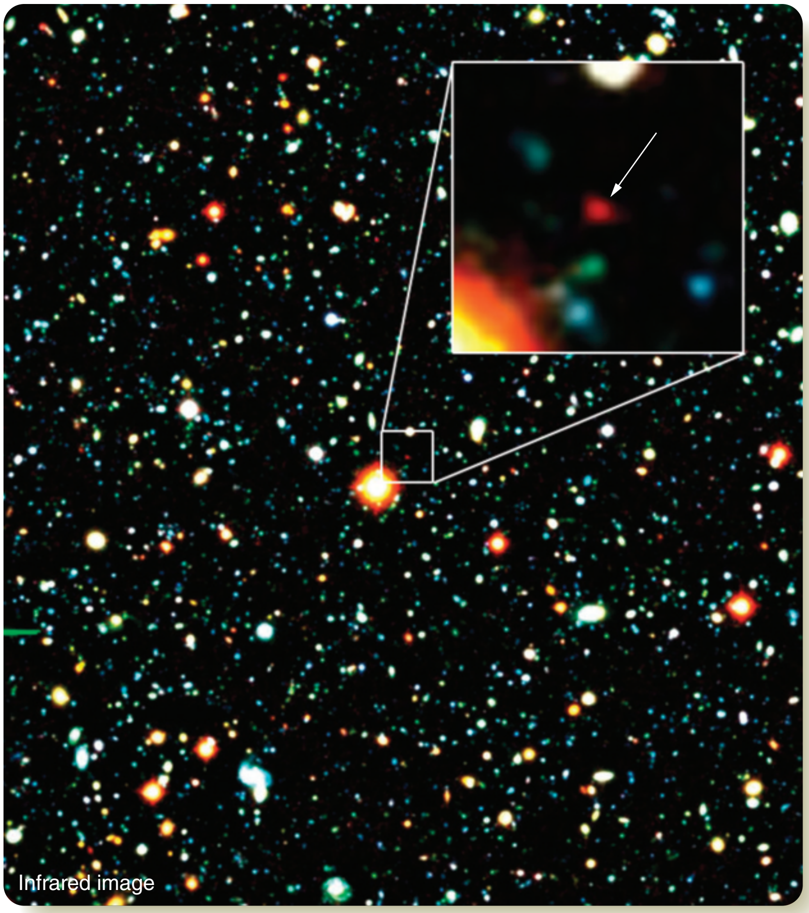


Figure 19-6

This faint galaxy is one of the most distant ever found. It has a redshift of 6.964, implying that it is 12.9 billion light-years from Earth. It appears as it was only 850 million years after the big bang when the light began its journey toward Earth.

occurred everywhere, and, in whatever direction you look, at great distance you can see back to the age when the universe was filled with hot gas (Figure 19-7).

The radiation that comes from such a great distance has a large redshift. The most distant visible objects are faint galaxies and quasars, with redshifts around 10. In contrast, the radiation from the hot gas right after the big bang has a redshift of about 1100. That means the light emitted as visible and near-infrared light by the hot gas in the early universe arrives at Earth as far-infrared, microwave, and radio waves. You can't see it with your eyes, but it can be detected with infrared and radio telescopes. Unlike the Gettysburg Address, the big bang can still be observed by the radiation it emitted. That amazing discovery is the subject of the next section.

The Cosmic Background Radiation

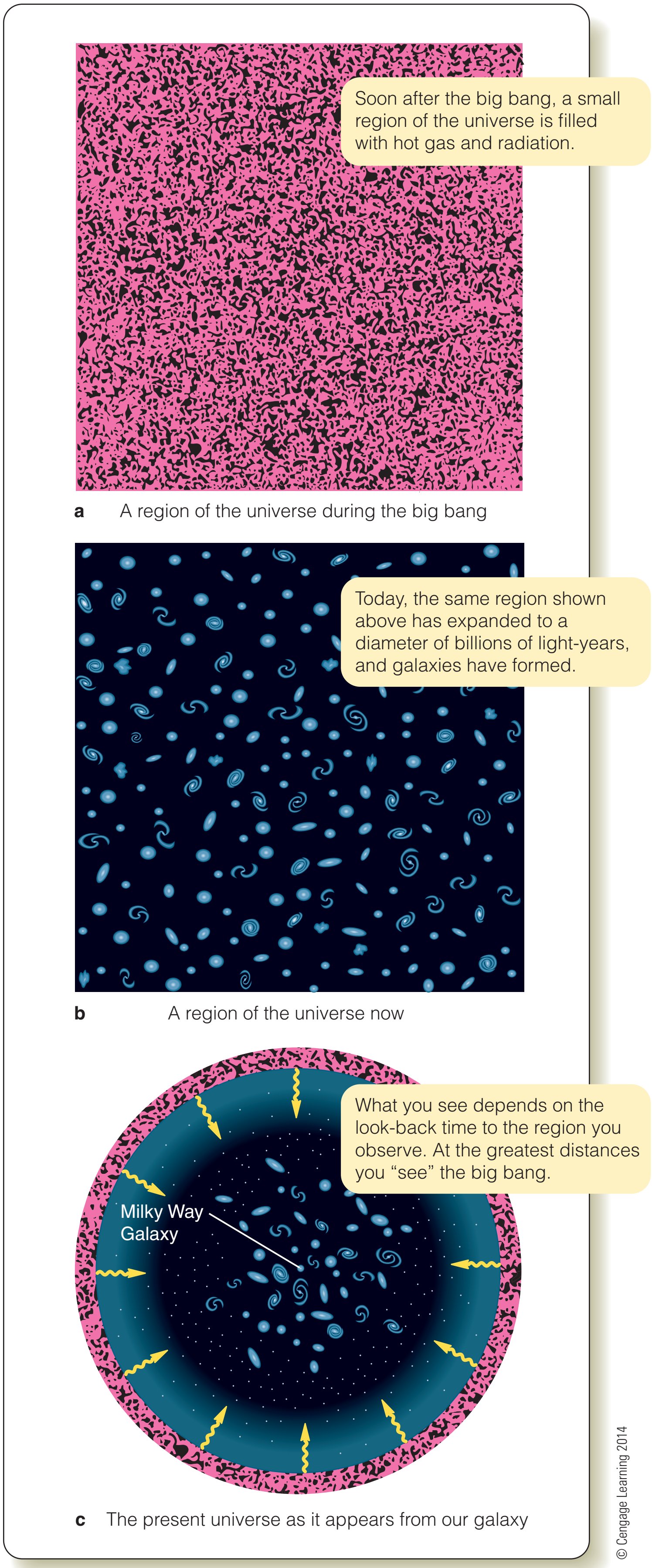
The story of the discovery of radiation from the time of the big bang begins in the mid-1960s when two Bell Laboratories physicists, Arno Penzias and Robert Wilson, were measuring the brightness of the sky at radio wavelengths (Figure 19-8). Their measurements showed a strange extra signal in the system, which they first attributed to the infrared glow from pigeon droppings inside the antenna. Perhaps they would have enjoyed scraping out the antenna more if they had known they would win the 1978 Nobel Prize in physics for the discovery they were about to make.

When the antenna was cleaned, they again measured the radio brightness of the sky and found the low-level noise was still there. The pigeons were innocent, but what was causing the extra signal?

The explanation for the noise goes back decades earlier. In 1939, astronomers noticed that spectra of some molecules in the interstellar medium showed they were bathed in radiation from a source with a temperature of 2 to 3 K. In 1948, physicist George Gamow predicted that the gases in the universe right after the big bang would have been hot and should have emitted strong blackbody radiation (see Chapter 6, pages 106–108). A year later, physicists Ralph Alpher and Robert Herman pointed out that the large redshift of the big bang material relative to Earth would lengthen the wavelengths of that radiation into the microwave part of the spectrum, with a blackbody temperature they estimated as 5 K. In the mid-1960s Robert Dicke at Princeton concluded the radiation should be just strong enough to detect with newly developed techniques. Dicke and his team began building a receiver. When Penzias and Wilson heard of Dicke's work, they recognized the mysterious extra signal they had detected as radiation from the big bang, the **cosmic microwave background** radiation.

The detection of the background radiation was tremendously exciting, but cosmologists wanted confirmation. Theory predicted that the radiation should have a spectrum like blackbody radiation coming from a very cool source, but the critical observations could not be made from the ground because Earth's atmosphere is opaque at the predicted blackbody peak wavelength. It was not until 1990 that satellite measurements confirmed the background radiation has exactly a blackbody spectral distribution, with an apparent temperature of 2.725 +/− 0.002 K—close to the original prediction.

It may seem strange that the gas of the big bang seems to have a temperature of just 2.7 degrees above absolute zero, but recall the tremendous redshift. Observers on Earth receive radiation that has a redshift of about 1100—that is, the wavelengths of the photons are about 1100 times longer than when they were emitted. The gas clouds that originally emitted the photons had a temperature of about 3000 K, and they emitted blackbody radiation with a λ_{max} of about 1000 nm (Wien's law, Chapter 6). Although that wavelength is in the near-infrared, the gas would also have emitted enough visible light to be seen glowing orange-red if there had been a human eye present at the time.



a A region of the universe during the big bang

b A region of the universe now

c The present universe as it appears from our galaxy

Figure 19-7

This diagram shows schematically the expansion of a small part of the universe. Although the universe is now filled with galaxies, the look-back time distorts what you see. Nearby you see galaxies, but at greater distances, the look-back time reveals the universe at earlier stages, before galaxies formed. At very great distances, the big bang is detectable as infrared, microwave, and radio energy arriving from the hot gas that filled the universe soon after the big bang.

The redshift has made the wavelengths of the background radiation about 1100 times longer on arrival at Earth than when they were emitted, so λ_{max} is now about 1 million nm, or 1 mm. That is why that hot gas seems to be about 2.7 K, 1100 times cooler than it actually was.

The first few steps up the cosmology pyramid have not been very difficult. Simple observations of the darkness of the night sky and the redshifts of the galaxies tell you that the universe must have had a beginning, and you have seen that the cosmic microwave background radiation is clear evidence of the big bang, that the early universe was hot and dense. Theorists can combine these observations with modern physics to add some details to the story of how the big bang occurred.

Photon and Particle Soup

Cosmologists cannot begin their history of the big bang at time zero, because no one understands the physics of matter and energy under such extreme conditions, but they can come surprisingly close. If you could visit the universe when it was only 10 millionths of a second old, you would find it filled with high-energy photons having a temperature well over 1 trillion (10^{12}) K and a density of 5×10^{18} kg/m^3, greater than the density of an atomic nucleus. (When cosmologists say the photons have a given temperature, they mean the photons have the same spectrum as blackbody radiation emitted by an object of that temperature.) Consequently, the photons in the early universe were gamma rays, with very short wavelength and therefore very high energy. (When cosmologists say that the radiation had a certain density, they refer to Einstein's equation $E = mc^2$. Using that equation, you can express a given amount of radiation energy per volume as if it were matter of a given density.)

If photons have enough energy, two photons can convert into a pair of particles—a particle of normal matter and a particle of **antimatter.** When an antimatter particle meets its matching particle of normal matter—when an antiproton meets a normal proton, for example—the two particles annihilate each other and convert their mass into energy in the form of two gamma rays. In the early universe, the photons were gamma rays and had enough energy to produce proton-antiproton pairs or neutron-antineutron pairs. When these particles collided with their antiparticles, they converted their mass back into photons. Thus, the early universe was filled with a dynamic soup of energy flickering from photons into particles and back again.

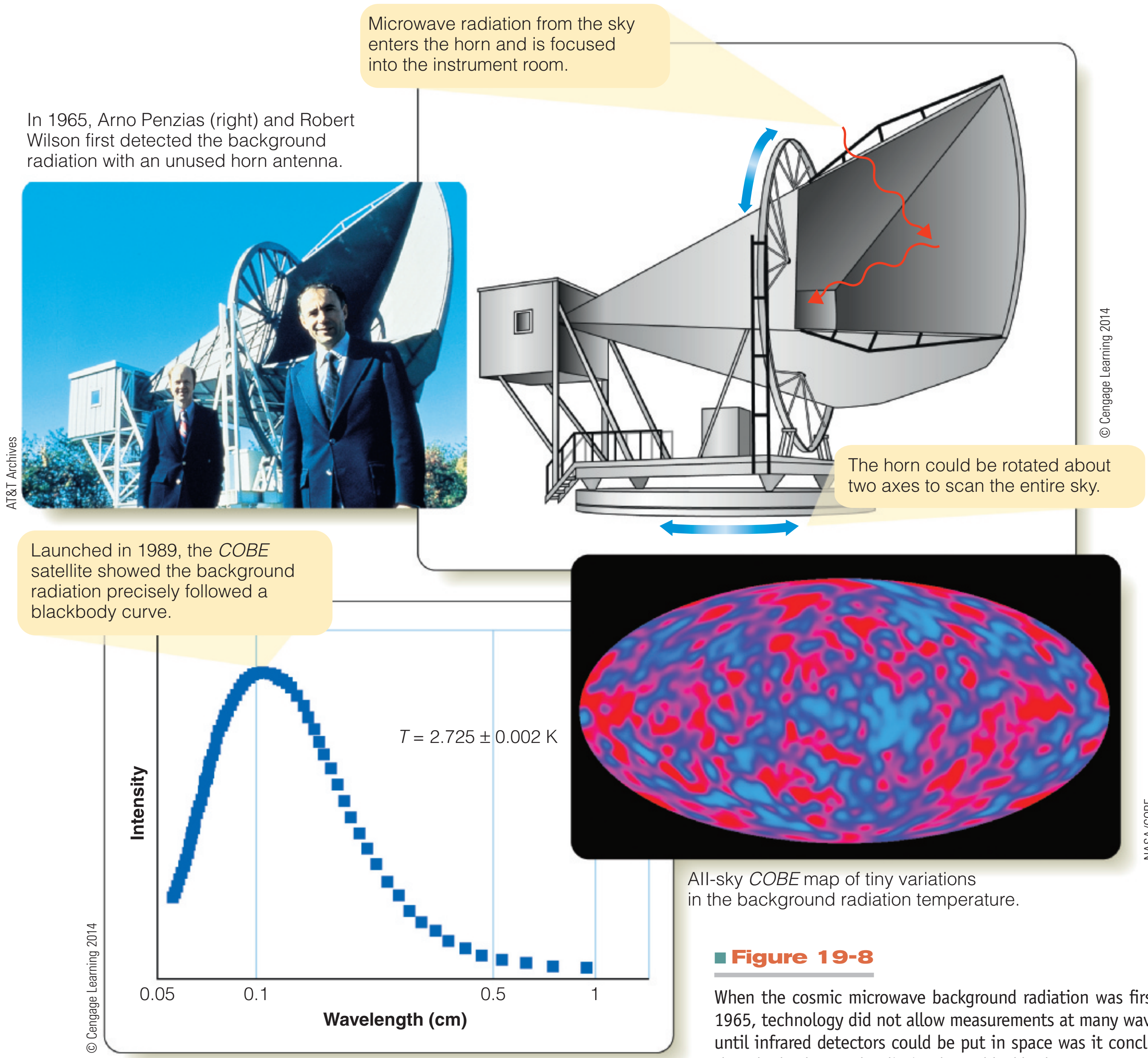


Figure 19-8

When the cosmic microwave background radiation was first detected in 1965, technology did not allow measurements at many wavelengths. Not until infrared detectors could be put in space was it conclusively shown that the background radiation has a blackbody spectrum, as predicted by theory.

While all this went on, the expansion of the universe caused the temperature of the radiation to drop, reducing the energy of the photons. When the universe was 0.0001 second old, its temperature had fallen to 10^{12} K. By that time, the average energy of the radiation photons had fallen below the energy equivalent to the mass of a proton or a neutron, so the gamma rays could no longer produce such heavy particles. Those particles that did exist combined with their antiparticles and quickly converted their mass into photons.

It would seem that all of the protons and neutrons should have been annihilated with their antiparticles; but, due to causes that are poorly understood, a small excess of normal particles existed. For every billion protons annihilated by antiprotons, one survived with no antiparticle to destroy it. Consequently, you live in a world of normal matter, and antimatter is very rare.

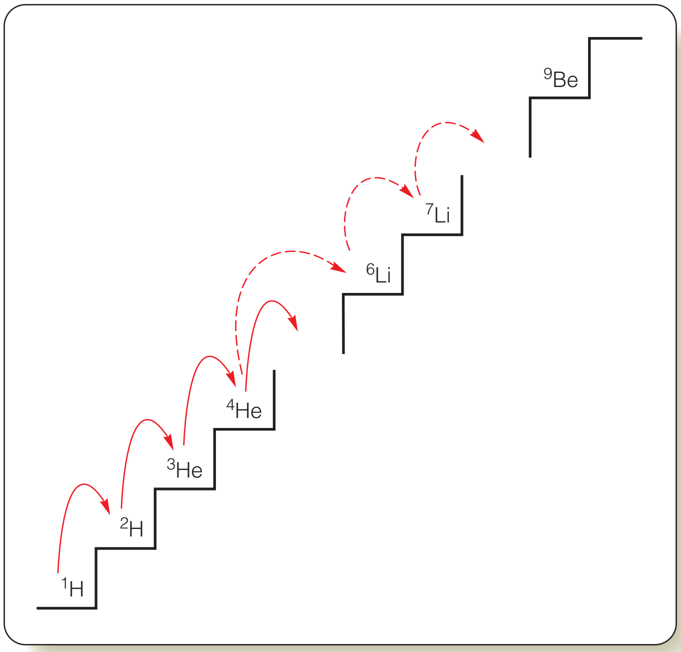
Although the gamma ray photons did not have enough energy when the universe was older than about 0.0001 second to produce any more protons and neutrons, they could still produce electrons and positrons (anti-electrons; look back to Chapter 7), which are about 1800 times less massive than protons and neutrons. That process continued until the universe was about 1 minute old, at which time the expansion had cooled to the point at which there were no remaining photons with enough energy to create electron-positron pairs. Again, most of the electrons and positrons combined to form photons, and only one in a billion electrons survived. Cosmologists can calculate, based on the known properties of subatomic particles and also the characteristics of the universe as a whole, that the populations of protons, neutrons, and electrons in our universe were produced during the first minute of its history.

A Few Minutes of Nucleosynthesis

The universal soup of hot gas and radiation continued to cool. Photons with high enough energy can break up atomic nuclei, so the formation of stable nuclei could not occur until the universe had cooled enough. By the time the universe was about 2 minutes old, protons and neutrons could link to form deuterium, the nucleus of a heavy hydrogen atom, without being immediately broken apart. By the end of the third minute, further reactions began converting deuterium into helium.

Almost no atoms heavier than helium could be built in the big bang, because there are no stable nuclei with atomic weights of 5 or 8 (in units of the hydrogen atom). Nuclei of atomic weights 5 and 8 are radioactive and decay almost instantly back into smaller nuclei. Cosmic element building during the big bang had to proceed step-by-step, like someone hopping up a flight of stairs (■ Figure 19-9). The lack of stable nuclei at atomic weights of 5 and 8 meant there were missing steps in the stairway, and the step-by-step reactions had great difficulty jumping over these gaps during the few minutes of the big bang. As a result, cosmologists can calculate that only a tiny amount of lithium (atomic weights 6 and 7) would have been produced during the big bang, and no heavier elements.

By the time it was 3 minutes old, the universe had become so cool that almost all nuclear reactions had stopped. By the time it was 30 minutes old, the nuclear reactions had ended completely, and about 75 percent of the mass of the universe was in the form of protons—hydrogen nuclei. The rest was helium nuclei. That composition, set during the first minutes of the universe, is the composition observed now for the oldest stars. Formation of elements with atomic weights greater than lithium had to wait for relatively slow-cooking nucleosynthesis processes in stars (look back to Chapter 14), beginning many millions of years after the big bang.



■ **Figure 19-9**

Cosmic element building: During the first few minutes of the big bang, temperatures and densities were high, and nuclear reactions built heavier elements. Because there are no stable nuclei with atomic weights of 5 or 8, the process built very few atoms heavier than helium.

Radiation and Matter; Recombination; and Re-ionization

After the era of nucleosynthesis, as the young universe continued to expand and cool, it went through a further sequence of four important changes. Originally, the universe was so hot that the gas was totally ionized, and the electrons were not attached to nuclei. The free electrons interacted with photons so easily that a photon could not travel very far before it encountered an electron and was deflected (■ Figure 19-10a). The photons interacted continuously with the matter, and the radiation and matter cooled together at a rate set by the expansion rate of the universe. In those circumstances, the universe was dominated by its radiation.

Cosmologists calculate that the first change occurred when the universe reached an age of about 50,000 years. Then, the density of energy in the form of photons became less than the density of the gas. Before that time, ordinary matter could not clump together because the intense sea of photons smoothed the gas out. As you will learn in the next section, dark matter, which does not interact with photons, was able to start clumping much earlier. After the 50,000-year mark, when the density of the universe's radiation permanently became less than that of its ordinary (baryonic) matter, ordinary matter could begin to draw together under the influence of the gravitational attraction of dark matter to form the clouds that eventually became galaxies.

The expansion of the universe spread the particles of the still-ionized gas farther and farther apart. When the universe reached an age of about 400,000 years, the second important change began. By then, the free electrons were spread so far apart that the photons could travel for thousands of parsecs before being deflected by an electron (Figure 19-10b). In other words, the universe started to become transparent. At approximately the same time, the third important change happened. As the falling temperature of the universe reached 3000 K, protons were able to capture and hold free electrons to form neutral hydrogen, a process called **recombination.** (This term is a little misleading, because the particles had never been together stably before; *combination* would be more accurate.) As the free electrons were gobbled up, the gas became essentially completely transparent, and the photons could travel through the gas without being deflected (Figure 19-10c).

After recombination, although the gas continued to cool, the photons no longer interacted with the gas, and consequently the photons retained the blackbody temperature that the gas had at recombination. Those photons, which started their journey with a blackbody temperature of 3000 K, are observed now as the cosmic microwave background radiation. Remember that the large

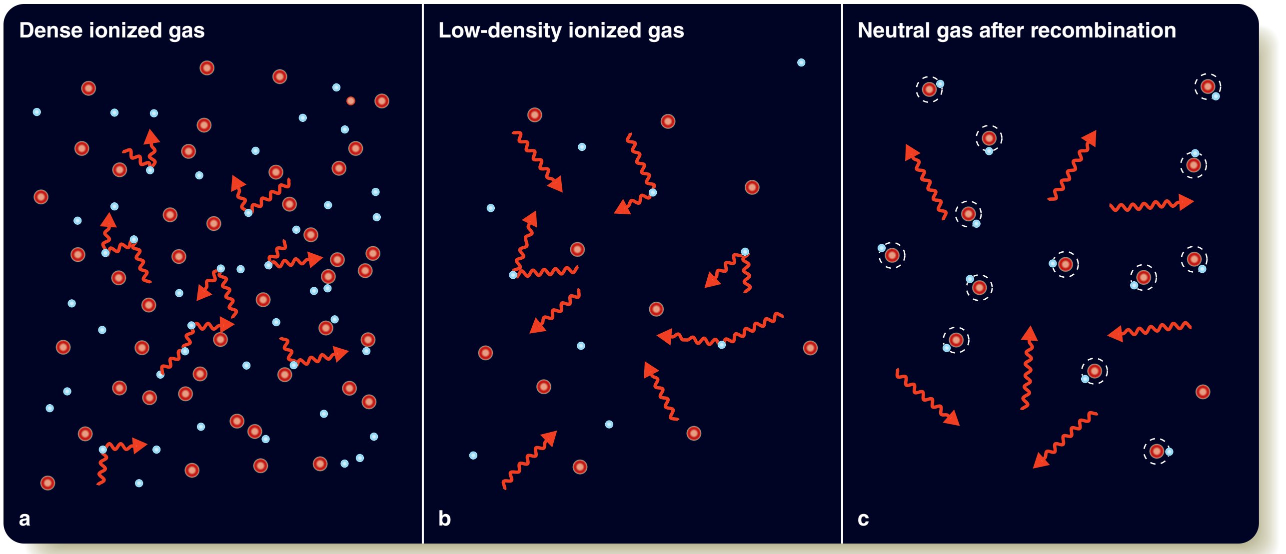


Figure 19-10

Photons scatter from electrons (blue) easily but hardly at all from the much more massive protons (red). (a) When the universe was very dense and ionized, photons could not travel very far before they scattered off an electron. This made the gas opaque. (b) As the universe expanded, the electrons were spread further apart, and the photons could travel farther; this made the gas more transparent. (c) After recombination, most electrons were locked to protons to form neutral atoms, and the gas became highly transparent.

redshift—the expansion of the universe—has stretched the wavelengths of the background radiation so that it appears now to have a temperature of about 2.7 K.

After recombination, the gas that originated in the big bang was neutral, warm, and transparent. As the universe continued to expand and cool, the glow from the warm gas faded into infrared wavelengths. The universe entered what cosmologists call the **dark age,** a period lasting about 400 million years until the formation of the first stars. During the dark age the universe expanded in darkness.

The fourth remarkable change happened as the first stars began to form, ending the dark age. The gas from which the first stars formed contained almost no metals and was therefore highly transparent. Mathematical models show that stars formed from this metal-poor gas would have been very massive, very luminous, and very short lived. That first violent burst of massive star formation produced enough ultraviolet light to begin ionizing the gas, and astronomers now, looking back to the most distant visible quasars and galaxies, can see traces of that **re-ionization** era in the universe (■ Figure 19-11). Re-ionization not only marked the end of the dark age, it was the beginning of the age of stars and galaxies in which you live now.

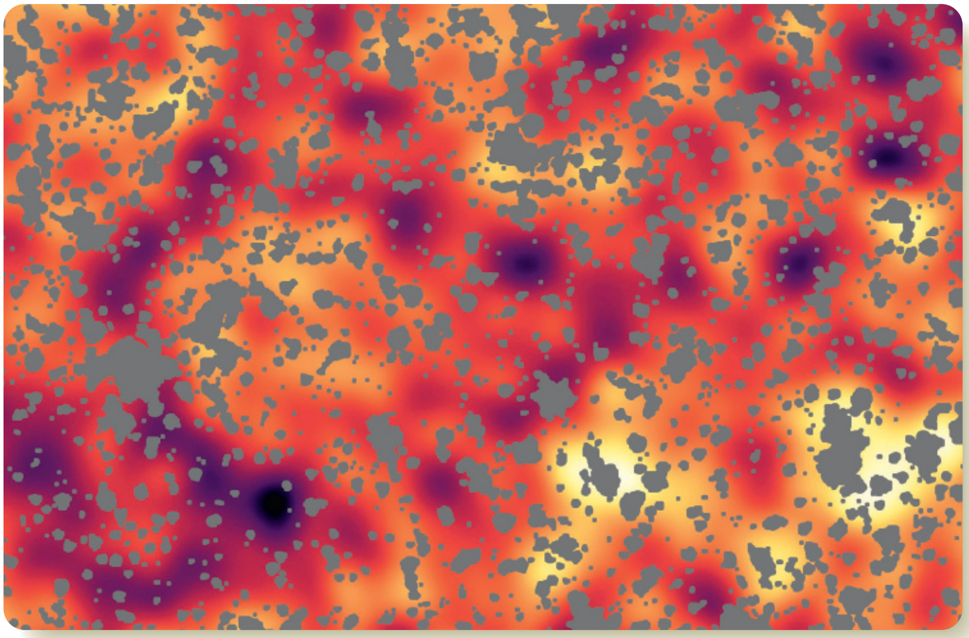
Look carefully at ■ Figure 19-12; it summarizes the story of the big bang, from the sorting out of stable particles in the earliest moments, through formation of helium in the first 3 minutes, up to the eras of energy-matter equality, recombination, and finally re-ionization of the gas. It may seem amazing that mere humans limited to Earth can draw such a diagram, but remember that it is based on evidence and on the best understanding of how matter and energy interact (How Do We Know? 19-2).

SCIENTIFIC ARGUMENT

How do you know there was a big bang?

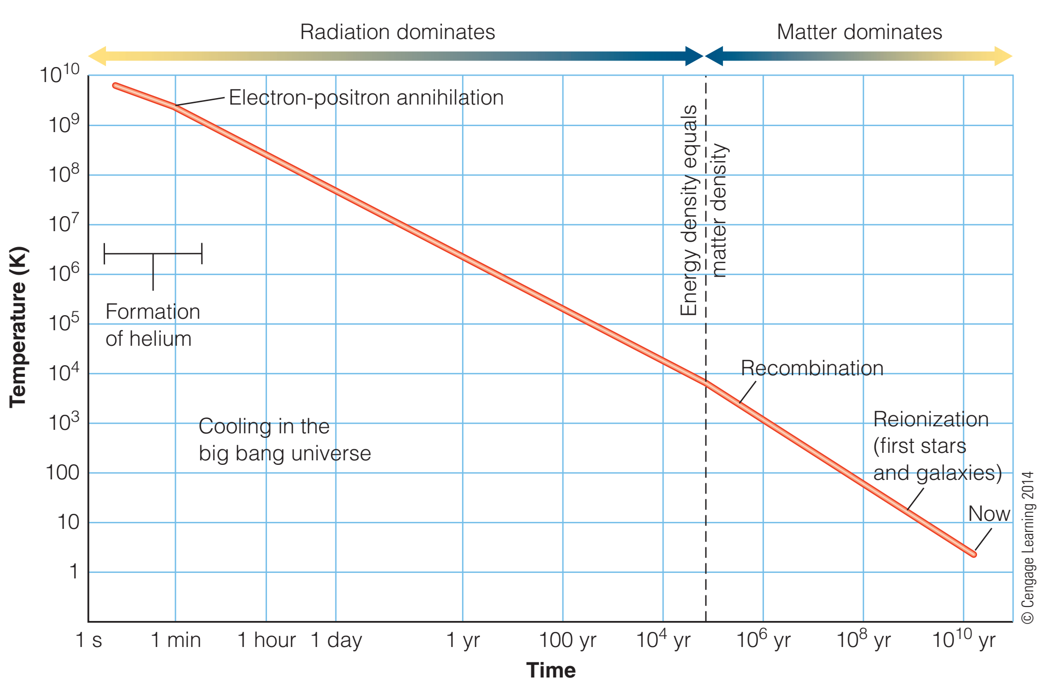
A good scientific argument combines evidence and theory to describe how nature works, and this question calls for a detailed argument. The cosmic microwave background radiation consists of photons emitted by warm gas after the big bang, so when astronomers detect those photons, they are in a sense "seeing" the big bang. Of course, all scientific evidence must be interpreted, so you need to understand how the big bang could produce radiation all around the sky before you can accept the background radiation as evidence. First, you must remember that the big bang event filled the entire universe with hot, dense gas. The big bang didn't happen in a single place; it happened everywhere. At recombination, the expansion of the universe reached the stage where the matter became transparent, and the radiation that had previously been trapped, bouncing between matter particles, was freed to travel through space. Now, that radiation from the age of recombination arrives from all over the sky. It is all around you because you are part of the big bang event, and as you look out into space to great distances, you look back in time and see the hot gas in any direction you look. You can't see the radiation as visible light because the large redshift has lengthened the wavelengths by a factor of 1100 or so, but you can detect the radiation as photons with infrared, microwave, and radio wavelengths.

With this interpretation, the cosmic microwave background radiation is powerful evidence that there was a big bang. That tells you how the universe began, but your argument includes an important point. **Why do you think the universe cannot have a center or an edge?**


NASA/JPL-Caltech/A. Kashlinsky (GSFC)

■ **Figure 19-11**

An infrared image from the *Spitzer Space Telescope* of a region in the constellation Ursa Major. Stars, galaxies, and other known sources have been removed by computer processing. The remaining background glow is from a period of time when the universe was less than one billion years old and is probably from the universe's very first objects such as massive stars and voracious black holes at the centers of forming galaxies. Yellow and white colors show the brightest parts of the background.



© Cengage Learning 2014

■ **Figure 19-12**

During the first few minutes of the big bang, some hydrogen was fused to produce helium, but the universe quickly became too cool for such fusion reactions to continue. The rate of cooling increased as matter began to dominate over radiation. Recombination freed the radiation from the influence of the gas, and re-ionization was caused by the birth of the first stars. Note how the exponential scale in time stretches early history and compresses recent history.

How Do We Know? | 19-2

Science: A System of Knowledge

What is the difference between believing in the big bang and understanding it? If you ask a scientist, "Do you believe in the big bang?" she or he may hesitate before responding. The question implies something incorrect about the way science works.

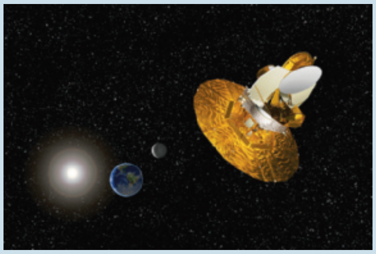
Science is a system of knowledge and process but not a system of belief. Science is an attempt to understand logically how nature works and is based on observations and experiments used to test and confirm hypotheses and theories. A scientist does not really believe in even a well-confirmed theory in the way people normally use the word *believe*. Rather, the scientist understands the theory and recognizes how different pieces of evidence support or contradict the theory.

There are other ways to know things, and there are many systems of belief. Religions, for example, are systems of belief that are not entirely based on observation. In some cases, a political system is also a system of belief; many people believe that democracy is the best form of government and do not ask for, or expect, evidence supporting that belief. A system of belief can be powerful and lead to deep insights, but it is different from science.

Scientists try to be careful with words, so thoughtful scientists would not say they *believe* in the big bang. They would say that the evidence is overwhelming that the big bang really did occur and that they are compelled by a logical analysis of both the observations and the theory to conclude that the theory is very likely correct. In this way scientists try to be objective and reason without distortion by personal feelings and prejudices.

A scientist once referred to "the terrible rule of evidence." Sometimes the evidence forces a scientist to a conclusion she or he does not like, but science is not a system of belief, so the personal preferences of each scientist must take second place to the rule of evidence.

Do you believe in the big bang? Or, instead, do you have confidence that the theory is right because of your analysis of the evidence? There is a big difference.



NASA / WMAP Science Team

Scientific knowledge is based objectively on evidence such as that gathered by spacecraft.

19-3 Space and Time, Matter and Energy

How CAN THE BIG BANG HAVE HAPPENED everywhere? To solve the puzzle, you need to put seemingly reasonable expectations aside and look carefully at how space and time behave on cosmic scales.

Looking at the Universe

The universe looks about the same in whichever direction you observe; in cosmologist language the universe is described as being **isotropic.** Of course, if you look toward a galaxy cluster you see more galaxies than in other directions, but that is only a local variation. On the average, you see similar numbers of galaxies in every direction. Furthermore, the background radiation is almost perfectly uniform across the sky. The universe is observed to be very isotropic.

The universe also seems about the same everywhere; in cosmologist language the universe is described as being **homogeneous.** (Note that this is a different property than the universe being isotropic, which means that it looks the same in all directions from your position.) Of course, there are local variations; some regions contain more galaxies and some fewer. Also, because the universe evolves, at large look-back times you see galaxies at an earlier developmental stage. But, if you account for those well-understood variations, then the universe seems to be, on average, the same everywhere.

The universe is observed to be isotropic and homogeneous. Those two properties don't have to go together; for an intellectual exercise, you can try to imagine arrangements that are isotropic but not homogeneous, or homogeneous but not isotropic. The fact that the universe is both isotropic and homogeneous leads to the **cosmological principle,** which says that any observer in any galaxy sees the same general properties for the universe. As you just learned, this overall principle intentionally overlooks minor local and evolutionary variations. The cosmological principle implies there are no special places in the universe. What you see from the Milky Way Galaxy is typical of what all intelligent creatures see from their respective home galaxies. Furthermore, the cosmological principle states, in a new form, the idea that the universe can have no center or edge. A center or an edge would be a special place, and the cosmological principle says there are no special places.

The Cosmic Redshift

Einstein's theory of general relativity, published in 1916, describes space and time together as the fabric of the universe, called space-time. That idea will give you a new insight into the meaning of the phrase *expanding universe.*

General relativity describes space-time as if it is made of stretching rubber, and that explains one of the most important

observations in cosmology—the redshifts. It is a **Common Misconception** that cosmological redshifts are Doppler shifts of galaxies flying away through space. Instead, except for relatively small, local motions within clusters of galaxies, galaxies are at rest. They are being *separated* from each other as space-time expands. Also, as space-time expands, it stretches any photon traveling through space so that its wavelength increases. Photons from distant galaxies spend more time traveling through space and are stretched more—in other words have larger redshifts—than photons from nearby galaxies. That is why redshift depends on distance.

Astronomers often express redshifts as if they were actual radial velocities, but the redshifts of the galaxies are not Doppler shifts. That is why this book is careful to refer to a galaxy's *apparent* velocity of recession. Some textbooks convert cosmological redshifts to velocities using Einstein's relativistic Doppler formula, but that formula applies to motion through space and not to the behavior of space-time itself, so that formula should not be used in a cosmological context. The Hubble law still applies: Redshifts can be used to find the distance to galaxies because redshifts show how much the universe has expanded, and thus how much time has elapsed, since the photons started on their journey.

It is common to think of space as just an emptiness, but now you have learned that space can have interesting properties of its own. Study **The Nature of Space-Time** on pages 438–439, and notice three important points.

1. Einstein described space-time as if it were a rubber canvas on which the universe is painted. The gradual stretching of space-time lengthens the wavelengths of photons on their way from distant galaxies to Earth. Thus the redshifts of the galaxies are not Doppler shifts.
2. Space-time could be curved. A model universe with positive curvature would be closed and finite. A zero-curvature model is said to be flat and must be open and infinite. A model with negative curvature is also open and infinite. In none of these models could the universe have an edge or a center.
3. For most of the 20th century, cosmologists attempted to measure the curvature of space-time. Modern observations indicate that the geometry of the universe is almost certainly flat and therefore the universe must be infinite.

Model Universes

Almost immediately after Einstein published his general relativity theory, other theorists were able to solve the sophisticated mathematics and compute simplified descriptions of the behavior of space-time and matter. The resulting model universes dominated cosmology throughout the 20th century.

The equations allowed three possibilities. Space-time might be curved in either of two different ways, or it might have no curvature at all. Most people find these curved models difficult to imagine, and modern observations have shown that the simplest model, without curvature, is almost certainly correct, so you don't have to wrap your brain around the curved models. You might, however, like to know a few of their most important properties.

Some models described space-time as being curved back on itself to form a **closed universe.** You would not notice this curvature in daily life; it would only be evident in measurements involving very distant objects. Throughout the 20th century, observations could not eliminate closed models, so they were considered a real possibility. Closed universe models have finite volumes, but, because space-time is curved back on itself, a closed universe nevertheless would have no edge and no center. Such closed models predicted that the expansion of the universe would eventually become a contraction that would bring all of the matter and energy back to the high-density big bang state, which was sometimes called the "big crunch."

Other models described an uncurved, **flat universe.** That is the kind of space-time you would expect from your daily life, with the same rules of geometry that you learned in high school, even for cosmological distances. Flat models are infinite, have no edge or center, and expand forever.

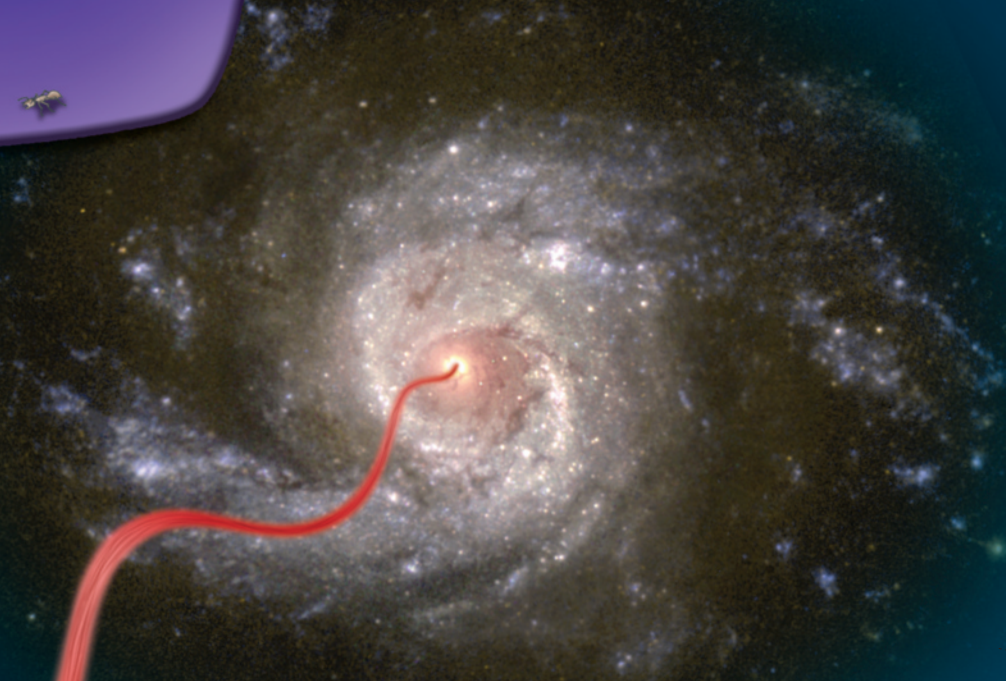
A third kind of model described an **open universe.** Such models contain curved space-time but are not curved back on themselves. Like the flat models, open models are infinite, have no center or edge, and expand forever.

Cosmologists of the 20th century struggled to find reasons or observations that would allow them to choose among these three kinds of models. The main clue was density. According to general relativity, the overall curvature of space-time is determined by the average density of matter plus energy in the universe. Cosmologists could calculate that space-time is flat if the average density of the universe equals a **critical density** of 9×10^{-27} kg/m^3. If the average density of the universe is more than the critical density, the universe must be closed; and if it is less, the universe must be open. Attempts to measure the actual value of the universe's density, however, were difficult and inconclusive.

These three kinds of universe models are illustrated in ■ Figure 19-13, which compares the expansions versus time of the different models. The parameter R on the vertical axis is a measure of the extent to which the universe has expanded. You can think of R, in a simplified way, as the average distance between galaxies. In the figure you can see that closed universes expand and then contract while both flat and open universes expand forever.

Notice that you can't find the actual age of the universe unless you know whether it is open, closed, or flat. The Hubble time is the age the universe would have if it were totally open, which would mean it contained almost no matter. If, instead, the universe is dense enough to have flat geometry, then its age

The Nature of Space-Time



NASA and The Hubble Heritage Team (STScI/AURA); Acknowledgment: G. R. Meurer and T.M. Heckman (JHU), and C. Leitherer, J. Harris and D. Calzetti (STScI), M. Sirianni (JHU)

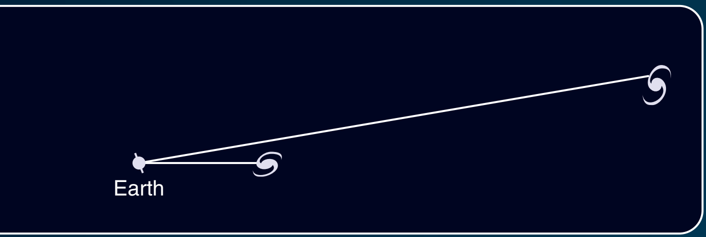
1 In 1929, Edwin Hubble discovered that the redshifts of the galaxies are proportional to distance—a relationship now known as the Hubble law. It was taken as dramatic evidence that the universe is expanding.

Distance is the separation between two points in space, and time is the separation between two events. Albert Einstein showed how to relate space to time and treat the whole as space-time. You can think of space-time as a canvas on which the universe is painted, a canvas that can stretch.

1a Astronomers often express galaxy redshifts as apparent velocities of recession, but these redshifts are not Doppler shifts. They are caused by the expansion of space-time.

For decades textbooks have described the cosmological redshifts using Einstein's relativisitic Doppler formula, but that formula applies to motion through space and not to the behavior of space itself.

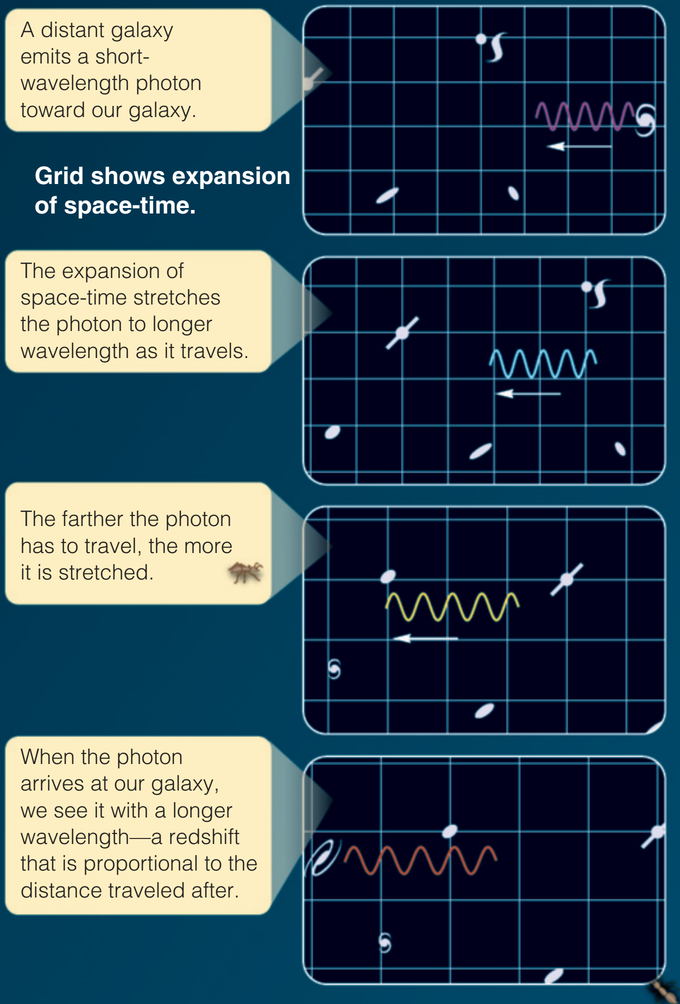
Galaxies are not really moving any more than the raisins in the raisin bread analogy are swimming through the bread dough. Except for orbital motion among neighbors, galaxies are motionless in space-time, and are being carried away from each other by the stretching of space-time.



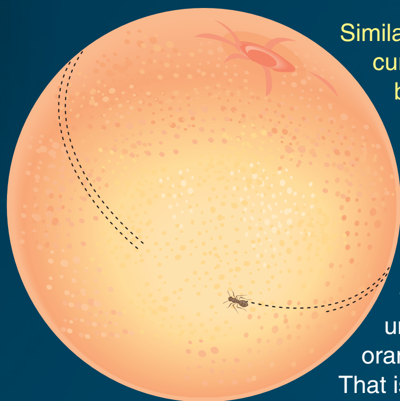
1c A photon traveling to Earth from a more distant galaxy spends more time moving through space and is stretched more than a photon from a nearby galaxy. That is why redshift is proportional to distance.

1b The stretching of space-time not only moves the galaxies away from each other, but it lengthens the wavelength of photons traveling through space-time. See illustration below.

What are the cosmological redshifts?

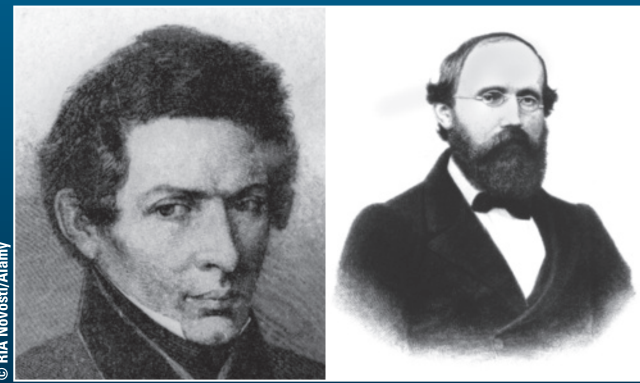


2 To think about space-time curvature, you can use an analogy of a two-dimensional ant on an orange. If he is truly two-dimensional, he will not be able to understand up and down. He can only travel forward and back, right and left. Then as he walks over the surface of his spherical universe, he will eventually realize he has been everywhere because his universe is covered by his footprints. He will conclude that he lives in a finite universe that has no edge and no center.



Similarly a model universe could be curved back on itself so that it could be finite but an explorer would never find an edge. With no edge, such a universe could have no center.

Notice that the center of the orange cannot be the center of the ant's two-dimensional universe because the center of the orange does not lie on the surface. That is, the center of the orange is not part of the ant's universe.



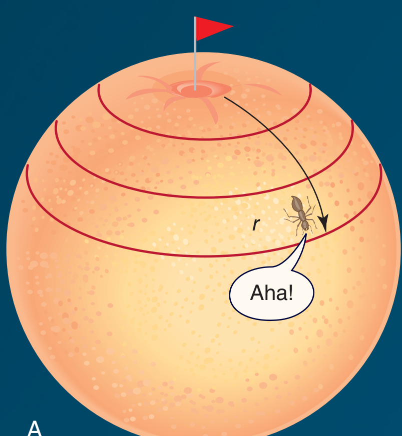
© RIA Novosti/Alamy

© INTERFOTO/Alamy

Nikolai Lobachevsky 1792–1856

Bernhard Riemann 1826–1866

During the 19th century, two mathematicians, the Russian Nikolai Lobachevsky and the German Bernhard Riemann, developed the mathematics of curved surfaces—non-Euclidian geometry. That, with Einstein's theory of general relativity, showed that space-time might be curved by an amount that depends on the average density of the universe. Curved space-time dominated 20th century cosmology.



2a There are three possible ways a model universe might be curved—positive, zero or negative. In your analogy, a positively curved universe could be represented by a sphere, a zero curvature universe by a flat surface, and a negatively curved universe by a saddle shape.

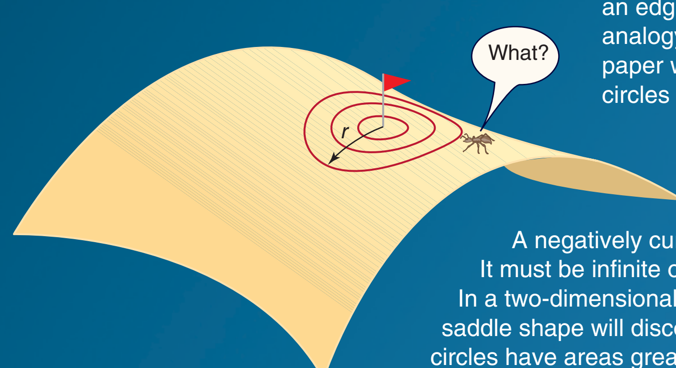
A positively curved model universe is termed closed because it curves back on itself. It is finite, but it has no edge. In a two-dimensional analogy, ants on an orange would notice that the areas of large circles were less than πr^2.

Don't forget: These analogies are only two-dimensional. When you think about our universe, you must think of a three-dimensional universe. It is as difficult for you to imagine curvature in our three-dimensional universe as it is for ants to imagine curvature in their two-dimensional universes.

A model universe with zero curvature is termed open because it does not curve back on itself. It must be infinite or it will have an edge. In a two-dimensional analogy, ants on a flat sheet of paper would find that all circles have areas of πr^2.

3 Throughout the 20th century, cosmologists tried to determine the curvature of our universe. Unlike the ants on this page, the cosmologists could not measure the areas of circles or the volumes of spheres, but they could measure the apparent velocities of recession of galaxies at great distances. Those measurements are difficult and were inconclusive. As you will see later in this chapter, just as the century ended, new ways to make measurements of the curvature became possible, and those show that the universe is flat.

If the universe truly has flat (Euclidean) geometry, then it must be infinite and has no center.



A negatively curved model is open. It must be infinite or it will have an edge. In a two-dimensional analogy, ants on a saddle shape will discover that larger circles have areas greater than πr^2.

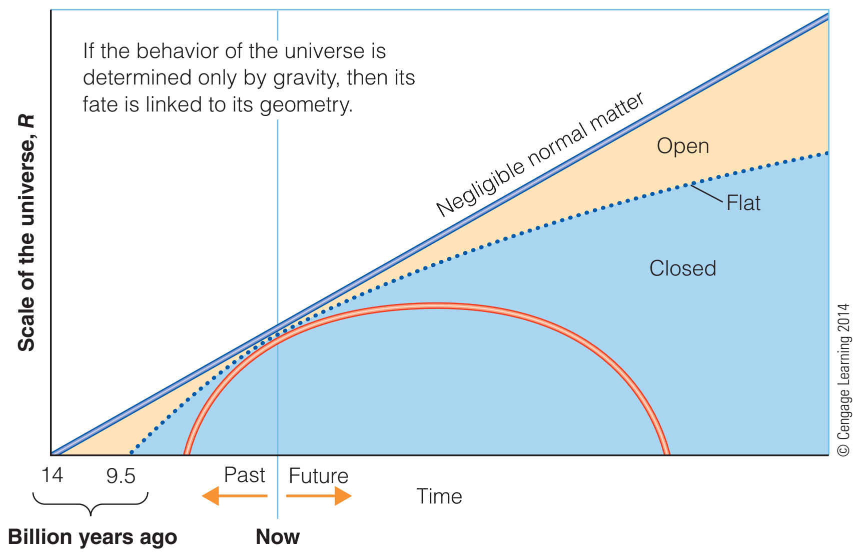


Figure 19-13

Illustrations of some simple universe models that depend only on the effect of gravity. Open universe models expand without end, and the corresponding curves fall in the region shaded orange. Closed models expand and then contract again (red curve). A flat universe (dotted line) marks the boundary between open and closed universe models. In these models the relationship between estimated age and actual age of the universe depends on the curvature of space-time. The age of the universe for each model is shown on the graph as the horizontal distance from "Now" back to the time when the universe scale factor R equaled 0.

would be ⅔ of the Hubble time. If H is about 70 km/sec/Mpc, as astronomers have measured, and the universe is flat, then it should be only 9.5 billion years old. Astronomers know of globular star clusters that are significantly older than that, so this was known as the age problem: How can the universe be younger than some of the objects in it?

Modern observations show that the universe is flat, and later in this chapter you will see how the evidence eliminated curved models and also solved the age problem. For now, however, you can think about a different puzzle. If the universe is flat, then its average density must exactly equal the critical density. Yet when cosmologists added up the matter that could be detected, they found only a few percent of the critical density. They wondered if the dark matter made up the rest.

Dark Matter in Cosmology

There is more to the universe than meets the eye. In Chapters 17 and 18, you discovered that the Milky Way Galaxy and other galaxies have much stronger gravitational fields than would be predicted based on the number of stars visible. Their gravitational fields are much stronger than expected even when you add in the gravity of the gas and dust that would be contained in ■ Figure 19-14. Gravitational lensing is dramatic evidence of dark matter. The galaxy cluster in the figure contains so much dark matter that it warps space-time and focuses the images of more distant galaxies into short arcs of light. The universe contains much more dark matter than normal matter. Looking at the universe of visible matter in galaxies is like looking at a tree and seeing only the leaves.

The available evidence shows that dark matter is not the same as the normal matter you are made of. To follow that argument, you need to estimate the density of normal matter, and you can start with what is known about conditions right after the big bang.

As you have learned, during the first few minutes of the big bang, nuclear reactions converted some protons into helium and a very small amount into other elements. How much of elements heavier than helium could be produced was controlled by the density of normal matter during the big bang. For example, if the density of matter had been high so that there were lots of normal particles such as protons and neutrons flying around, they would have collided with the deuterium nuclei and converted most of them into helium. On the other hand, if the density of normal particles had been relatively low, more deuterium would have survived intact without being converted to helium.

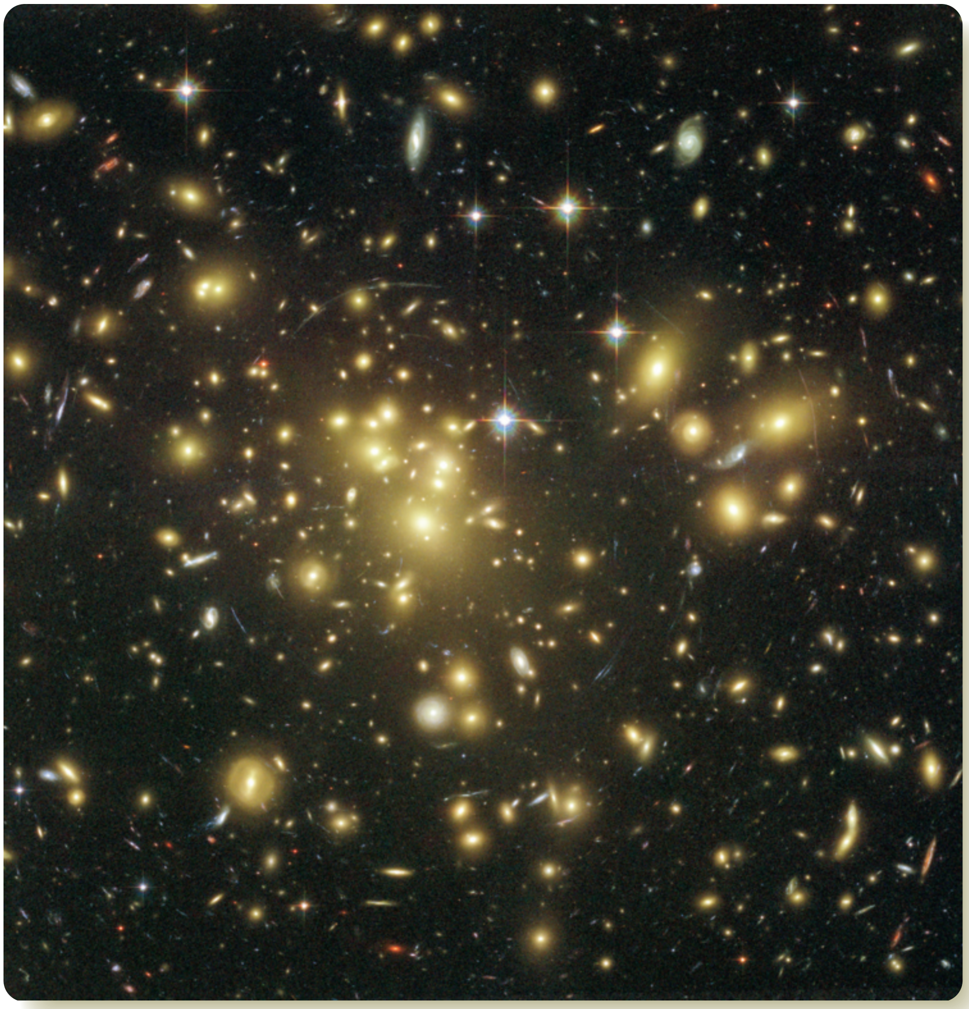
Figure 19-9 shows that there is a gap between helium and lithium; there is no stable nucleus with atomic mass 5, so regular nuclear reactions during the first few minutes of the big bang had difficulty converting helium into lithium. If, however, the density of protons and neutrons was high enough, a few nuclear reactions could have leaped the gap and produced a small amount of the isotope lithium-7.

Both deuterium and lithium are destroyed in stars, so astronomers have attempted the difficult task of studying clouds of gas at large look-back times that had not yet been altered by nuclear reactions in stars. As shown in ■ Figure 19-15, the amount of deuterium observed now in the universe sets a lower limit on the density of the universe relative to the critical density, and the observed abundance of lithium-7 sets an upper limit.

The surprising conclusion from these observations and calculations is that normal matter from which you, Earth, and the stars are made cannot make up much more than 5 percent of the critical density.

The protons and neutrons that make up normal matter belong to a family of subatomic particles called *baryons,* so cosmologists believe that most of the dark matter cannot be baryons. Only a few percent of the mass in the universe can be ordinary baryonic matter; the dark matter must be **nonbaryonic** matter.

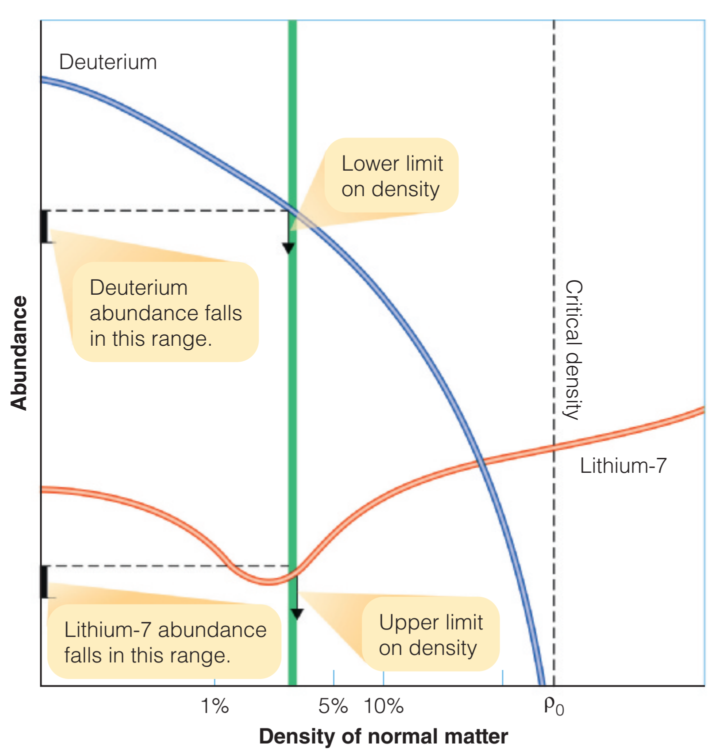
Some theorists thought that neutrinos, particles that are predicted to be very abundant in the universe yet are not baryons, might have enough mass to make up the dark matter, but modern measurements show that neutrinos are not massive



NASA, ESA, L. Bradley [Johns Hopkins University], R. Bouwens [University of California, Santa Cruz], H. Ford [Johns Hopkins University], and G. Illingworth [University of California, Santa Cruz]

Figure 19-14

Gravitational lensing shows that galaxy clusters contain much more mass than what is visible. The yellowish galaxies in this image are members of a relatively nearby cluster of galaxies. Most of the objects in this image are blue or red images of very distant galaxies focused by the gravitational field of the cluster. Some of these imaged galaxies may be over 13 billion light-years away. That they can be seen at all is evidence that the foreground galaxy cluster contains large amounts of dark matter.



© Cengage Learning 2014

Figure 19-15

This diagram compares observation with theory. Theory predicts how much deuterium (blue curve) and lithium-7 (red curve) you would observe for different densities of normal matter. The observed density of deuterium falls in a narrow range shown at upper left and sets a lower limit on the possible density of normal matter. The observed density of lithium-7, shown at lower left, sets an upper limit. This means the true density of normal matter must fall in a narrow range represented by the green column. Certainly, the density of normal matter is much less than the critical density.

enough. Some particle physics theories predict the existence of a new type of particle that are labeled **WIMPs** (weakly interacting massive particles), but they have not been detected with certainty in the laboratory. The true nature of dark matter remains one of the major mysteries of astronomy.

Theoretical models of galaxy formation in the early universe provide more hints about the nature of dark matter. Dark matter composed of particles moving at or near the speed of light is referred to as **hot dark matter.** Such fast-moving particles would not clump together easily and could not have stimulated the formation of objects as small as galaxies and clusters of galaxies. Instead, the most successful models of galaxy formation require that the dark matter actually be made up of **cold dark matter,** meaning the particles move slowly and can clump into relatively small structures.

Nonbaryonic dark matter does not interact significantly with normal matter or with photons, which is why you can't see it. This also means that dark matter was not affected by the intense radiation that dominated the universe when it was very young. So long as radiation dominated the universe, it prevented clouds of normal matter from contracting to begin forming galaxies. But the dark matter was immune to the radiation, and the dark matter could contract to form clumps while the universe was very young. In other words, dark matter could have given galaxy formation a head start soon after the big bang. Models with cold, nonbaryonic dark matter are most successful at predicting the formation of galaxies and clusters of galaxies with the right size and at the right time in the history of the universe to match the observations.

Although the evidence is very strong that plenty of dark matter exists, it is still not abundant enough to make the universe flat. Dark matter constitutes less than 30 percent of the critical density. As you will see later in this chapter, there is more to the universe than meets the eye, more even than dark matter.

SCIENTIFIC ARGUMENT

Why do cosmologists think that dark matter can't be baryonic?
Good scientific arguments always fall back on evidence. In this case the evidence is very compelling. Small amounts of isotopes like deuterium and lithium-7 were produced by nuclear reactions in the first minutes of the big bang, and the resulting abundance of those elements depends strongly on how many protons and neutrons were available at that time. Because those particles belong to the family of particles called baryons, physicists refer to normal matter as baryonic. Measurements of the abundances of deuterium and lithium-7 show that the universe cannot contain more baryons than about 5 percent of the critical density. Yet, observations of galaxies and galaxy clusters show that dark matter must make up almost 30 percent of the critical density. Consequently, cosmologists conclude that the dark matter must be made up of nonbaryonic particles.

Finding the dark matter is important because the density of matter in the universe determines the curvature of space-time. Now build a new argument. **How does the modern understanding of space-time explain cosmic redshifts?**

19-4 Twenty-First-Century Cosmology

If you are a little dizzy from the weirdness of expanding space-time and dark matter, make sure you are sitting down before you read further. As the 21st century began, astronomers made a discovery that startled all cosmologists: The expansion of the universe is actually accelerating. You'll have to go back a couple of decades to understand how that amazing new discovery nevertheless fits well with some of the things you have learned already in this chapter.

Inflation

By 1980, accumulating evidence had made the big bang theory widely accepted by cosmologists, but it faced two problems that led to the development of a new theory—a revised big bang with an important addition.

One of the problems is called the **flatness problem.** The properties of the universe appear to be close to the dividing line between being open or closed; that is, the geometry of the universe seems nearly flat, and therefore its density is close to the critical density. Given the vast range of possibilities, from zero to infinity, it seems peculiar that the density of the universe is close to the critical value. If dark matter is as common as the measurements indicate, the universe's density must be within a factor of 3 of the critical density.

Even a small departure from critical density when the universe was young would be magnified by subsequent expansion. To be so near critical density now, the density of the universe during its first moments must have been within 1 part in 10^{49} of the critical density. So the flatness problem can be stated as: Why is the universe so close to perfectly flat?

The second problem with the original big bang theory is called the **horizon problem.** This is related to the observed isotropy of the cosmic microwave background radiation. When astronomers make a correction for the motion of Earth, they see the same intensity and temperature of background radiation in all directions to a precision of better than 1 part in 1000. Yet when you observe background radiation coming from two points in the sky separated by an angle larger than about 1 degree, you are looking at two bits of matter that seemingly should never have been connected from the time of the big bang up to the time when the radiation was emitted. In other words, when recombination occurred and the gas of the big bang became transparent to the background radiation, the universe was not old enough for any form of energy to have traveled from one of those two regions to the other. Thus, according to the standard big bang theory, the two spots you observe had not had enough time, since the start of the big bang, to exchange energy and equalize their temperatures. (The term *horizon* is used in this context

because the two spots are said to lie beyond their respective light-travel horizons.) Then, how did every part of the entire big bang universe get to be so nearly the same temperature at the time of recombination?

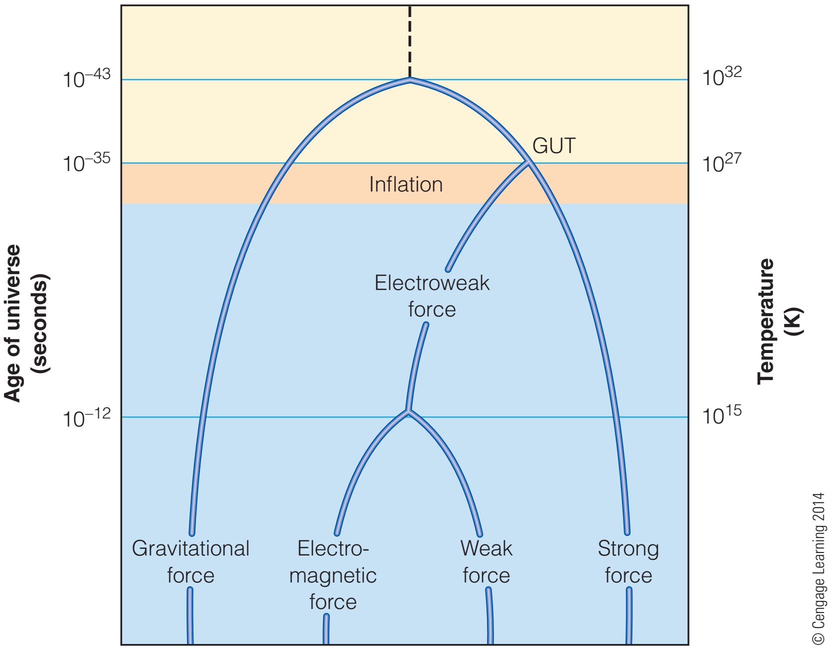
The key to these two problems and to others involving subatomic physics may lie with the hypothesis called the **inflationary universe.** That hypothesis predicts there was a sudden enormous expansion called *inflation* when the universe was very young, even more extreme than that predicted by the original big bang theory.

To understand the inflationary universe, you need to recall that physicists know of only four forces—gravity, the electromagnetic force, the strong force, and the weak force (look back to Chapter 7, page 131). You are familiar with gravity; you fought it to get out of bed this morning. The electromagnetic force is responsible for making magnets stick to refrigerator doors and cat hair stick to sweaters charged with static electricity, as well as holding electrons in orbit around atomic nuclei and being intimately connected with processes that make light and radiation. The strong force holds atomic nuclei together, and the weak force is involved in certain kinds of radioactive decay.

For many years, theorists have tried to unify these forces; that is, they have tried to describe all four forces as aspects of a single mathematical law. A century ago, James Clerk Maxwell showed that the electric force and the magnetic force were really the same effect, and physicists now count them as a single electromagnetic force. Similarly, in the 1960s, theorists succeeded in unifying the electromagnetic force and the weak force in what is now called the electroweak force. Those two forces operate effectively as a single unified force, but only in very high-energy processes. At lower energies the electromagnetic force and the weak force behave differently. Now, theorists have proposed ways of unifying the electroweak force with the strong force at even higher energies. These new theories are called **grand unified theories,** or **GUTs.**

According to the inflationary universe hypothesis, the universe expanded and cooled until about 10^{-36} seconds after the big bang, when it became cool enough that the electroweak force and the strong force began to disconnect from each other; that is, they began to behave in different ways (■ Figure 19-16). Cosmologists calculate that this change would have released tremendous energy, which then suddenly inflated the universe by a size factor of 10^{50} or larger. At the start of inflation, the part of the universe that is now observable from Earth is estimated to have been a factor of 10^{35} smaller than a proton, but it suddenly inflated to roughly a meter across and then continued its slower expansion to its present extent.

Early rapid inflation of the universe can solve both the flatness problem and the horizon problem. The sudden inflation of the universe would have forced whatever curvature it had before that moment toward a value of zero, just as inflating a balloon makes regions on its surface flatter. Consequently, you now live in a universe that is almost perfectly flat because of that long-ago moment of inflation. In addition, because the entire observable part of the universe was only about 10^{-59} light seconds across when the universe was about 10^{-36} seconds old, it would have had enough time to equalize its temperature before inflation occurred. As a result, you now live in a universe in which the background radiation has the same temperature in all directions.



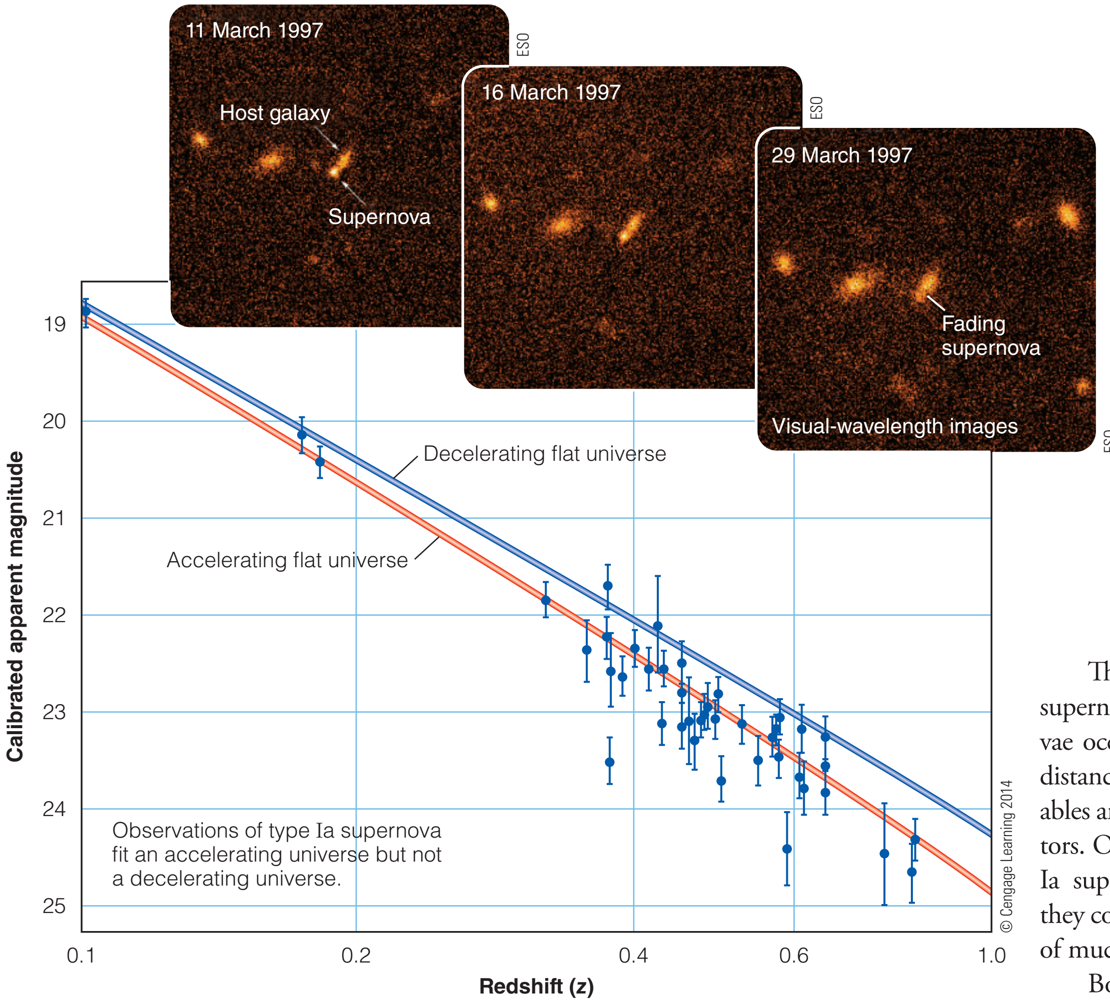
■ **Figure 19-16**

When the universe was very young and hot (top), the four forces of nature were indistinguishable in behavior. As the universe began to expand and cool, the forces "separated," meaning that they began to have different characteristics, which released a huge amount of energy and triggered a sudden rapid inflation in the size of the universe.

The inflationary theory predicts that the universe is almost perfectly flat, which means that the actual density equals the critical density. A theory can never be used as evidence, of course, but the beauty and simplicity of the inflationary theory have given many cosmologists confidence that the universe must be flat. Observations, however, seem to say that the universe does not contain enough matter (ordinary baryonic matter plus nonbaryonic dark matter) to be flat. Can there be more to the universe than baryonic matter and dark matter? What could be weirder and more obscure than dark matter? Read on.

The Acceleration of the Universe

Both common sense and the theory of general relativity suggest that as the galaxies recede from each other, the expansion should be slowed by gravity trying to pull the galaxies toward each other. How much the expansion is slowed should depend on the amount of matter in the universe. If the density of matter and energy in the



■ **Figure 19-17**

From the way supernovae fade over time, astronomers can identify those that are type Ia. Once calibrated, the peak brightness of each of those supernovae could be compared with their respective redshifts, revealing that distant type Ia supernovae were about 25 percent fainter than expected. That must mean they are farther away than expected, given their redshifts. This is strong evidence that the expansion of the universe is accelerating.

universe is less than the critical density, the expansion should be slowing only somewhat, and the universe should expand forever. If the density of matter and energy is greater than the critical density, the expansion should be slowing down dramatically, and the universe should eventually begin contracting. Notice that this is the same as saying a low-density universe should be open and a high-enough-density universe should be closed.

For decades, astronomers struggled to measure the redshifts and distances to very distant galaxies and detect the slowing of the expansion. Detecting a change in the rate of expansion is difficult because it requires accurate measurements of the distances to very remote galaxies.

The launch of the Hubble Space Telescope in 1990 made it possible to measure distances to galaxies with unprecedented accuracy, and two competing research teams began using the same technique of calibrating type Ia supernovae as distance indicators. As you learned in Chapter 15, a type Ia supernova occurs when a white dwarf gains matter from a companion star, exceeds the Chandrasekhar limit, and collapses in a supernova explosion. Because all such white dwarfs should collapse at the same mass threshold, they should all produce explosions of the same size and luminosity, which makes them good distance indicators.

The two teams calibrated type Ia supernovae by locating such supernovae occurring in nearby galaxies, with distances known from Cepheid variables and other reliable distance indicators. Once the peak luminosity of type Ia supernovae had been determined, they could be used to find the distances of much more distant galaxies.

Both teams announced their results in 1998. They agree that the expansion of the universe is not slowing down. Contrary to expectations, it is speeding up! That is, the expansion of the universe is accelerating (■ Figure 19-17).

The announcement that the expansion of the universe is accelerating was totally unexpected, and astronomers immediately began testing it. This result depends critically on the calibration of type Ia supernovae as distance indicators, and some astronomers suggested that the calibration might be wrong (look back to How Do We Know 17-1). Problems with the calibration have been ruled out subsequently by supernovae observed at even greater distances. The universe really does seem to be expanding faster and faster rather than slowing down.

Dark Energy and Acceleration

If the expansion of the universe is accelerating, then there must be a force of repulsion in the universe that counteracts gravity, and cosmologists are struggling to understand what it could be. One possibility leads back to 1916.

When Albert Einstein published his theory of general relativity in 1916, he recognized that his equations describing spacetime implied that the galaxies could not float unmoving in space

because their gravity would pull them toward each other. The only solutions seemed to be either a universe that was contracting under the influence of gravity or a universe in which the galaxies were rushing away from each other so rapidly that gravity could not pull them together. In 1916, cosmologists did not yet know that the universe was expanding, so Einstein made what he later thought was a mistake.

To balance the attractive force of gravity, Einstein added a constant to his equations called the **cosmological constant,** represented by an uppercase lambda **(Λ).** That constant represents a force of repulsion that balances the gravitational attraction between galaxies so the universe would not contract or expand. Thirteen years later, in 1929, Edwin Hubble announced his observations indicating that the universe is expanding, and Einstein said introducing the cosmological constant was his biggest blunder. Modern cosmologists think he may have been right after all.

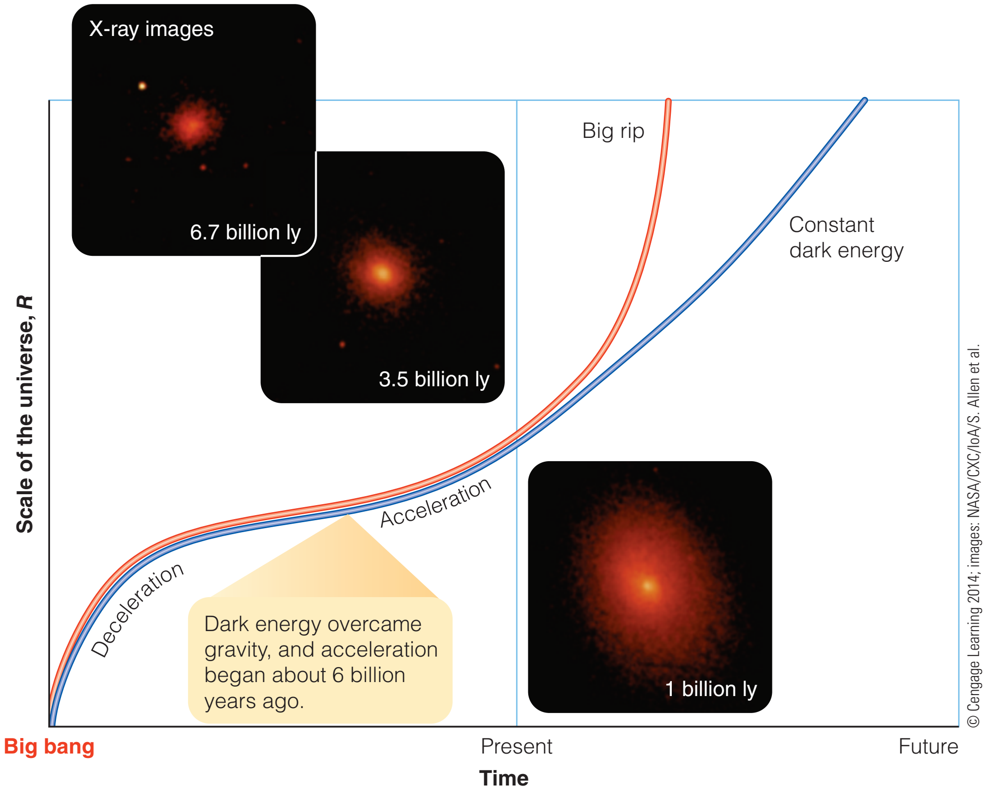
One explanation for the acceleration of the universe is that there is, after all, a cosmological constant, representing a type of antigravity force that is part of the fabric of space-time and causes a continuing acceleration in the expansion of the universe. Because the cosmological constant by definition remains constant with time, the universe would have experienced this acceleration throughout its history.

Another solution is to suppose that totally empty space, the vacuum, contains energy that drives the acceleration. This is an interesting possibility because theoretical physicists have long discussed the idea of a vacuum energy for reasons based on the behavior of subatomic particles. Cosmologists label a universal vacuum energy as **quintessence.** Unlike the cosmological constant, the effect of quintessence would not necessarily remain constant over time.

Whatever the correct explanation might be, whether a cosmological constant antigravity force, or a universal vacuum energy, or something else, the observed acceleration of the universe's expansion is evidence that a type of unknown energy is spread throughout space. Cosmologists refer to it as **dark energy,** an energy that drives the acceleration of the universe but does not contribute to the formation of starlight or the cosmic microwave background radiation.

You will recall that acceleration and dark energy were first discovered when astronomers found that supernovae a few billion light-years away were slightly fainter than expected based on their redshifts. The acceleration of the expansion has made those supernovae a bit farther away than implied by their redshifts, and so they look fainter. Since that discovery, astronomers have continued to find even more distant type Ia supernovae, some as distant as 12 billion light-years. The more distant of those supernovae are not too faint; they are too bright! That reveals even more about dark energy (■ Figure 19-18).

Unlike the medium-distant supernovae, the very distant supernovae are a bit too bright because they are not as far away as expected, and that confirms a theoretical prediction based on dark energy. When the universe was young, galaxies were closer together so that their gravitational pull on each other was stronger than the effect of dark energy and slowed the expansion. That makes the very distant supernovae a bit too bright. As space-time expanded, it moved the galaxies farther apart, their gravitational pull on each other became weaker than dark energy, and acceleration began. That makes the medium-distant supernovae a bit too faint. In other words, the observations show that sometime about 6 billion years (equal to 45 percent of the age of the universe) ago, the universe



■ **Figure 19-18**

X-ray observations of hot gas in galaxy clusters confirm that in its early history the universe was decelerating because gravity was stronger than the dark energy. As expansion weakened the influence of gravity, dark energy began to cause acceleration. The evidence is not conclusive, but it most directly supports the cosmological constant form of dark energy and weighs against quintessence, which means the universe may not undergo a "big rip." This diagram is only schematic, and the two curves are drawn separated for clarity; at the present time, the two curves have not diverged from each other.

shifted gears from deceleration to acceleration. The calibration of type Ia supernovae allows astronomers to observe this change from deceleration to acceleration.

Furthermore, dark energy can help you understand the curvature of the universe. The theory of inflation makes the specific prediction that the universe is flat. Recent observations you will learn about in the rest of this chapter confirm that prediction. Dark energy seems to explain how the universe can have enough matter plus energy to be flat. As you have already learned, $E = mc^2$ means that energy and matter are equivalent. Thus, the dark energy is equivalent to an amount of mass spread through space. Baryonic matter plus dark matter makes up about one-third of the critical density, and dark energy appears to make up the remaining two-thirds. That is, when you include dark energy, the total density of the universe equals the critical density, which makes the universe flat.

Step by step, you have been climbing the cosmological pyramid. Each step has been small and logical, but look where it has led you. You now know some of nature's deepest secrets, but you can imagine there are yet more steps above yet to be found, and more secrets to explore.

The Age and Fate of the Universe

Acceleration helps with another problem. The Hubble constant equals 70 km/s/Mpc, and earlier in this chapter you calculated the Hubble time, an estimate of the age of the universe, and found it was about 14 billion years. Further, you calculated a more precise estimate of the universe's age on the assumption that its geometry is flat, which gives an age of two-thirds of the Hubble time, or ⅔ times 14 billion years, which equals about 9 billion years. That was a puzzle because the globular star clusters are older than that. Now you are really ready to solve the age problem and then consider the fate of the universe.

If the expansion of the universe has been accelerating, then it must have been expanding more slowly in the past, and the galaxies would have taken longer to arrive at the average separations you can observe now. That means the universe can be flat but nevertheless older than two-thirds of the Hubble time. The latest estimates suggest that acceleration makes the age of a flat universe almost 14 billion years, coincidentally about the same as the simple Hubble time estimate. That age is clearly older than the oldest known star clusters, which solves the age problem.

For many years, cosmologists have enjoyed saying, "Geometry is destiny." Thinking of open, closed, and flat universe models, they concluded that the density of a model universe determines its geometry, and its geometry determines its fate. By this they meant that if the universe is open it must therefore expand forever, whereas if it is closed, it must eventually begin contracting. But that is true only if the behavior of the universe as a whole is ruled completely by gravity. If dark energy causes an acceleration that can dominate gravity, then geometry is *not* destiny, and, depending on the precise properties of dark energy, even a closed universe might expand forever.

The ultimate fate of the universe depends on the nature of dark energy. If dark energy is described by the cosmological constant, then the force driving acceleration does not change with time, and our flat universe will expand forever with the galaxies getting farther and farther apart and using up their gas and dust making stars, and stars dying, until each galaxy is isolated, burnt out, dark, and alone. If, however, dark energy is described by quintessence, then the force may be increasing with time, and the universe might accelerate faster and faster as space pulls the galaxies away from each other, eventually pulling the galaxies apart, then pulling the stars apart, and finally tearing individual atoms apart. This possibility has been called the **big rip.**

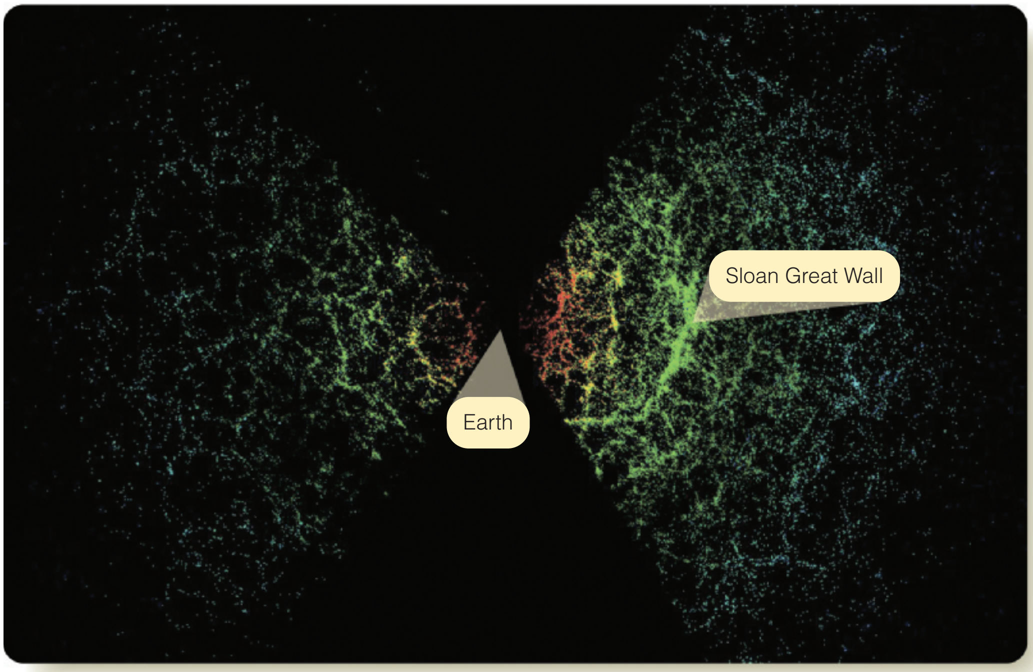
Probably there will be no big rip; critically important observations made by the Chandra X-Ray Observatory have been used to measure the amount of hot gas and dark matter in almost 30 galaxy clusters. Because the distance of each cluster is one of the variables in the calculations, the X-ray astronomers could compare the amount of hot gas and dark matter and solve for distance. The farthest of the clusters is 8 billion light-years away. These observations are important for two reasons. First, the redshifts and distance of these galaxies independently confirm the conclusion from the supernova observations that the expansion of the universe first decelerated but changed to accelerating about 6 billion years ago. Second, the Chandra results almost completely rule out quintessence. If dark energy is described by the cosmological constant and not by quintessence, then there will be no big rip (Figure 19-18).

The Origin of Large-Scale Structure

On the largest scales, the universe is isotropic. That is, it looks the same in all directions. But, on smaller scales, there are irregularities. The sky is filled with galaxies and clusters of galaxies that seem to be related to their neighbors in even larger aggregations that astronomers call large-scale structure. Studies of **large-scale structure** lead to insights about how the universe has evolved.

When you look at galaxies in the sky you see them in clusters ranging from a few galaxies to thousands, and those clusters appear to be grouped into **superclusters.** The Local Supercluster, in which you live, is a roughly disk-shaped swarm of galaxy clusters 50 to 75 Mpc in diameter. By measuring the redshifts and positions of hundreds of thousands of galaxies in great slices across the sky, astronomers have been able to create maps revealing that superclusters are not scattered at random. They are distributed in long, narrow **filaments** and thin walls that outline great **voids** nearly empty of galaxies (■ Figure 19-19).

This large-scale structure is a problem because the cosmic microwave background radiation is very uniform, and that means the gas of the big bang must have been extremely uniform at the time of recombination. Yet the look-back time to the



■ **Figure 19-19**

Nearly 70,000 galaxies are plotted in this double slice of the universe extending outward in the plane of Earth's equator. The nearest galaxies are shown in red and the more distant in green and blue. The galaxies form filaments and walls enclosing nearly empty voids. The Sloan Great Wall, discovered by the *Sloan Digital Sky Survey* (look back to Chapter 5), is almost 1.4 billion light-years long and is the largest known structure in the universe. The most distant galaxies in this diagram are roughly 3 billion light-years from Earth.

furthest galaxies and quasars is about 93 percent of the way back to the big bang. How did the uniform gas at the time of recombination become lumpy and coagulate so quickly to form galaxies? How did it make clusters of galaxies and the supermassive black holes we observe as quasars (look back to Chapter 18) so early in the history of the universe?

Baryonic matter is so rare in the universe that cosmologists can calculate that it did not have enough gravity to pull itself together quickly after the big bang. As you learned earlier, cosmologists propose that dark matter is nonbaryonic and therefore immune to the intense radiation that prevented normal matter from contracting. Dark matter, in contrast, was able to collapse into clouds and then attract normal matter to begin the formation of galaxies, clusters, and superclusters. Mathematical models have been made to describe this process, and cold dark matter does seem capable of jump-starting the formation of structure of the right sizes and in the right amount of time (■ Figure 19-20).

But what started the clumping of the dark matter? Theorists say that space is filled with tiny, random quantum mechanical fluctuations smaller than the smallest atomic particles. At the moment of inflation, those tiny fluctuations would have been stretched to very large but very subtle variations in gravitational fields that could have stimulated the formation of galaxy superclusters, filaments, and walls. The structure you see in Figure 19-19 may be the present-day ghostly traces of microscopic random fluctuations in the infant universe.

The Two-Degree-Field Redshift Survey mapped the position and redshift of 250,000 galaxies and 30,000 quasars. As expected, the galaxies are spread in filaments and walls, and a statistical analysis of the distribution matches the prediction of inflation and also shows that the universe expansion is accelerating. This is an important result because it is yet another confirmation of acceleration that is independent of the brightness of type Ia supernovae or X-ray–emitting gas in galaxy clusters. When a hypothesis is confirmed by observations of many different types, scientists have much more confidence that it is a true description of nature.

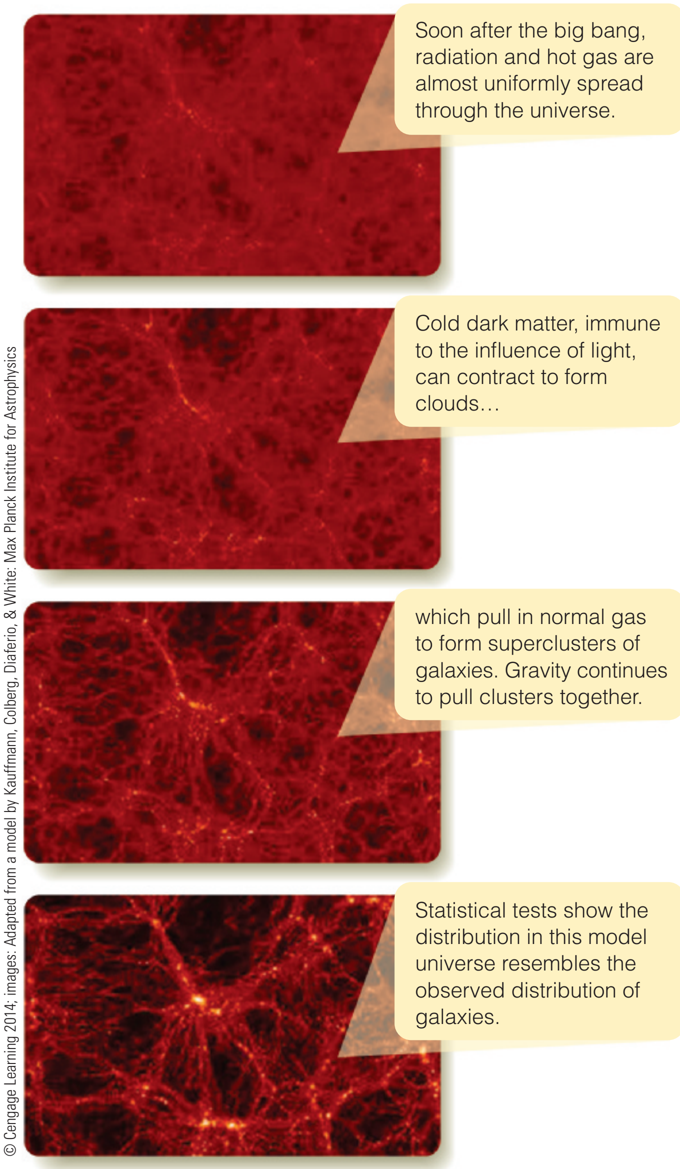
Curvature of the Universe

Astronomers continue to search for more ways to confirm that the universe is flat and that its expansion is accelerating. One way to measure the curvature of space-time is to measure the size of things at great look-back times, for example, by observing irregularities in the background radiation. That type of measurement is yielding impressive results.

The background radiation is nearly isotropic—it looks almost exactly the same in all directions after effects of Earth's motion and also emission by material in the foreground are taken into account. After those corrections are made to a map of the background radiation and the average intensity is subtracted from each spot on the sky, minor irregularities are evident. That is, some spots on the sky look a tiny bit hotter and brighter, or cooler and fainter, than the average (look back at Figure 19-8). Those irregularities contain lots of information.

The inflation-modified big bang theory of the universe makes specific predictions about the angular sizes of the variations you should find in the background radiation. Observations made by the COBE satellite in 1992 detected the variations with the largest diameter, but detecting smaller variations is critical for testing the theory. Six teams of astronomers have built specialized telescopes to make these observations. Some have flown under balloons high in the atmosphere, and others have observed from the ice of Antarctica. Extensive observations of the background radiation have been made by the *Wilkinson Microwave*

Growth of Structure in the Universe



© Cengage Learning 2014; images: Adapted from a model by Kauffmann, Colberg, Diaferio, & White: Max Planck Institute for Astrophysics

Figure 19-20

This computer model traces the formation of structure in the universe from soon after the big bang to the present.

Anisotropy Probe (WMAP). In 2009 the Planck space telescope was launched to do an even more thorough job.

It is a **Common Misconception** that explosions in space produce sound. Science-fiction movies imply that sound can travel through a vacuum and that exploding space ships make big *kabooms*. That's not true now, of course, but the early universe was dense enough that sound could travel through the gas, and the big bang did make a noise. A theorist described it as "a descending scream, building to a deep rasping roar, and ending in a deafening hiss." The pitch of the sound was about 50 octaves too low for you to hear, but those powerful sound waves did have an effect on the universe. They determined the size of the irregularities now detectable in the cosmic microwave background radiation.

Theory predicts that most of the irregularities in the hot gas of the big bang should be about one degree in diameter if the universe is flat. If the universe were open, the most common irregularities would be smaller. Careful measurements of the size of the irregularities in the cosmic background radiation show that the observations fit the theory very well for a flat universe, as you can see in ■ Figure 19-21. Not only is the theory of inflation confirmed, an exciting result itself, but these data show that the universe is flat, meaning space-time has no overall curvature, which indirectly confirms the existence of dark energy and the acceleration of the universe.

Cosmologists can analyze those microwave background irregularities using sophisticated mathematics to find out how frequently spots of different angular sizes occur. The calculations confirm that spots about 1 degree in diameter are the most common, but spots of other sizes occur as well, and it is possible to plot a graph such as ■ Figure 19-22 to show how frequently different sizes of irregularities occur.

The data points from the *WMAP* observations follow a wiggling line in Figure 19-22, and the size and positions of those wiggles tell cosmologists a great deal about the universe. The details of the curve shows that universe is flat, accelerating, and will expand forever. The age of the universe derived from the data is 13.7 billion years. Furthermore, the smaller peaks in the curve reveal that the universe contains 4.6 percent baryonic (normal) matter, 23.3 percent nonbaryonic dark matter, and 72.1 percent dark energy. The Hubble constant is confirmed to be 71 km/s/Mpc. The inflationary theory is confirmed, and the data give more support to the cosmological constant version of dark energy, although quintessence is not quite ruled out. Hot dark matter is ruled out. The dark matter needs to be cold dark matter to clump together rapidly enough after the big bang to make the galaxy clusters and superclusters we observe. In fact, a model based on these data predicts that the first stars began to produce light when the universe was only about 400 million years old.

Please reread the preceding paragraph. Especially to people who have been working in the field for years, that collection of firm facts is mind-blastingly amazing. *WMAP* and other studies of the cosmic microwave background radiation and the distribution of galaxies have revolutionized cosmology. At last, astronomers have accurate observations against which to test theories. The values of the basic cosmic parameters are known to a precision of 1 percent or better.

On reviewing these results, one cosmologist announced that "Cosmology is solved!", but that might be premature. Cosmologists

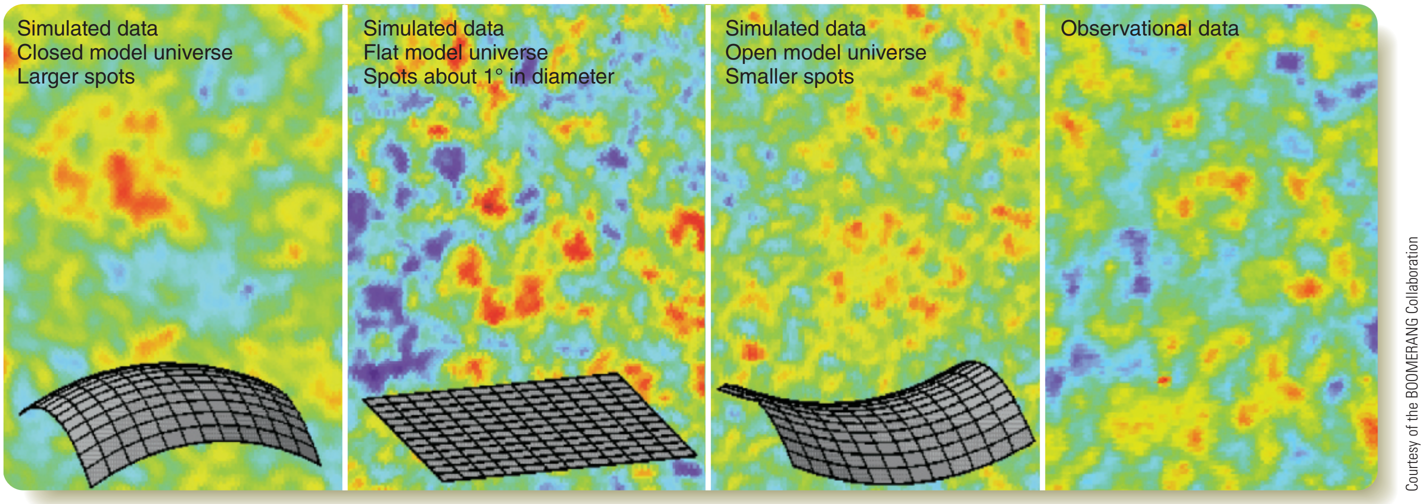


Figure 19-21

You can see the difference yourself. Compare the observations of the irregularities in the background radiation in the far right panel with the three simulations starting from the left. The observed size of the irregularities fits best with cosmological models having flat geometry. Detailed mathematical analysis confirms your visual impression: The universe is flat.

don't understand dark matter or the dark energy that drives the acceleration, so in a sense, over 95 percent of the universe is not understood. Hearing this, another cosmologist suggested a better phrase was "Cosmology in crisis!" Certainly there are further mysteries to be explored, but cosmologists are growing more confident that they can describe the origin and evolution of the universe.

SCIENTIFIC ARGUMENT

How does inflation theory solve the flatness problem?

This question requires a carefully constructed argument. The flatness problem can be stated as a question: Why is the universe so flat? After all, the density of matter in the universe could be anything from zero to infinite, but the observed sizes of variations in the cosmic background radiation show that the universe is flat, and therefore the average density of the universe is very close to the critical density. Furthermore, the density must have been astonishingly close to the critical density when the universe was very young, or it would not be as close as it is now. The inflationary theory solves this problem by proposing that the universe had a moment of rapid inflation when it was a tiny fraction of a second old. That inflation drove the universe toward flatness just as blowing up a balloon makes a spot on the balloon flatter and flatter.

Understanding theory in cosmology is critically important, but science depends on evidence. Now build another argument based on evidence. **What evidence can you cite that the expansion of the universe is accelerating?**

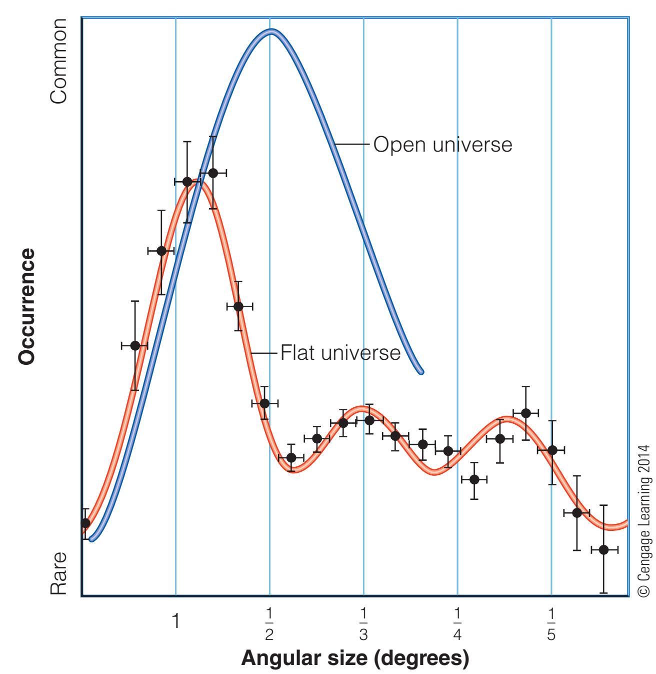


Figure 19-22

This graph shows how commonly irregularities of different sizes occur in the cosmic microwave background radiation. Irregularities of about 1 degree in diameter are most common. Models of the universe that are open or closed are ruled out. The data fit a flat model of the universe very well. Crosses on data points show the uncertainty in the measurements.

What Are We? Products

As you climbed the cosmology pyramid, you negotiated many steps. Most were easy, and all were logical. They have carried you up to some sweeping views, and you have traced the origin of the universe, the formation of the chemical elements, the birth of galaxies, and the births and deaths of stars. You now have a perspective that few humans share.

We are among the products of two cosmic processes, gravitational contraction and nuclear fusion. Gravity created instabilities in the hot matter of the big bang and triggered the formation of clusters of galaxies. Further contraction of that material formed individual galaxies and then stars within the galaxies. As stars began to shine through the universe, nuclear fusion in their cores began to cook hydrogen and helium into the heavier atoms from which humans are made.

As you review the history and structure of the universe, it is wise to recognize the mysteries that remain; but note that they are mysteries that may be solved rather than mysteries that are unknowable. Only a century ago, humanity didn't know there were other galaxies, or that the universe was expanding, or that stars generate energy by nuclear fusion. Human curiosity has solved many of the mysteries of cosmology and will solve more during your lifetime.

Study and Review

Summary

- **Cosmology (p. 425)** is the study of the structure and history of the universe as a whole. Astronomers and physicists who do research in cosmology are called **cosmologists (p. 426).**
- Cosmologists conclude that it is impossible for the universe to have an edge because an edge would introduce logical inconsistencies. In other words, an edge to the universe does not make sense. If the universe has no edge, then it cannot have a center.
- The darkness of the night sky leads to the conclusion that the universe is not infinitely old. If the universe were infinite in extent, infinite in age, and **static (p. 426),** then every spot on the sky would glow as brightly as the surface of a star. This problem, commonly labeled **Olbers's paradox (p. 426),** implies that the universe had a beginning.
- The **observable universe (p. 428),** the part you can see from Earth with a large telescope, is limited in size. It may be only a tiny portion of the entire universe, which could be infinite.
- Edwin Hubble's 1929 discovery that the redshift of a galaxy is proportional to its distance, now known as the Hubble law, shows that the galaxies appear to be moving away from each other. That phenomenon is called the **expanding universe (p. 428).**
- Tracing the expansion of the universe backward in time brings you to imagine an initial high-density, high-temperature state commonly called the **big bang (p. 429).**
- A rough estimate of the age of the universe based on the presently observed expansion rate is called the **Hubble time (p. 429).**
- Although the expanding universe began from the big bang, it has no center. The galaxies do not move away from a single point. Cosmologists understand that galaxies remain approximately in their respective positions and are carried further away from each other as the space between them expands.
- The **cosmic microwave background (p. 430)** radiation is blackbody radiation with a temperature of about 2.73 K, spread nearly uniformly over the entire sky. This radiation is the light from the big bang, freed from the gas at the time of **recombination (p. 433)** and now redshifted by a factor of 1100.
- The background radiation is clear evidence that the universe began with a big bang.
- During the earliest moments of the universe, matter and **antimatter (p. 431)** particles continually flashed in and out of existence. A slight excess of ordinary matter remained after most of the matter and antimatter particles annihilated each other.
- During the first few minutes of the big bang, nuclear fusion converted some of the hydrogen into helium but was unable to make many other heavy atoms because no stable nuclei exist with weights of 5 or 8. Now, hydrogen and helium are common in the universe, but heavier atoms are rare.
- After recombination, for a period of hundreds of millions of years called the **dark age (p. 434),** the universe expanded in darkness until the first stars came into existence. Astronomers have observed signs of **re-ionization (p. 434)** of the universe caused by that first generation of stars.
- The chemical composition of the oldest stars is about 75 percent hydrogen and 25 percent helium, which is what models of the big bang nuclear processes would predict. This is further evidence supporting the big bang theory.
- The universe is **isotropic (p. 436)** and **homogeneous (p. 436).** In other words, in its major features, the universe looks the same in all directions from Earth and also appears to have the same properties in all locations.
- Isotropy and homogeneity lead to the **cosmological principle (p. 436),** the idea that there are no special places in the universe. Except for minor local differences, every place is the same, and the view from every place is the same.
- Einstein's theory of general relativity explains that cosmic redshifts are caused by wavelengths of photons stretching as they travel through expanding space-time.
- **Closed universe (p. 437)** models are finite in size, but their space-time is curved back on itself so they have no edge or center.
- **Flat universe (p. 437)** models have uncurved space-time and are infinite. Modern observations show that the universe is probably flat.
- **Open universe (p. 437)** models have curved space-time, but it is not curved back on itself. Such universes are infinite.
- Open model universes contain less than the **critical density (p. 437)** and closed model universes more. If the universe is flat, then its density must equal the critical density.
- The amounts of deuterium and lithium-7 in the universe shows that normal baryonic matter can make up only about 5 percent of the critical density. Dark matter must be **nonbaryonic (p. 440)** and makes up less than 30 percent of the critical density.
- One suggestion for what dark matter might be is a hypothetical type of subatomic particle labeled **WIMPs (p. 442).** The observed ranges of sizes and masses for galaxy clusters disagrees with models that assume rapidly moving **hot dark matter (p. 442)** particles that would not clump together easily, and instead agrees with models that assume slowly moving readily clumping **cold dark matter (p. 442).**
- The **inflationary universe (p. 443),** a modification to the big bang theory, proposes that the universe briefly expanded dramatically, just a tiny fraction of a second after the big bang.
- The energy to drive inflation would have been released when the four forces of nature changed their respective properties to become different as the universe cooled in its earliest moments. This "separation" of forces is predicted by **grand unified theories (GUTs) (p. 443)** that explain the forces of nature as being aspects of a single force that are unified only in particle interactions with very high energies.
- Inflation explains the **flatness problem (p. 442)** because the large expansion forced the universe to become flat, just as a spot on an inflating balloon becomes flatter as the balloon inflates.
- Inflation explains the **horizon problem (p. 442)** because what is now the observable part of the universe was so small before inflation that energy could move and equalize the temperature everywhere in that volume.
- Observations of type Ia supernovae reveal the surprising fact that the expansion of the universe is speeding up. Cosmologists propose that this acceleration is caused by energy present in empty space labeled **"dark energy" (p. 445).**
- The nature of dark energy is unknown. It may be described by Einstein's **cosmological constant (Λ) (p. 445),** or its strength may change with time, which cosmologists label **"quintessence" (p. 445).** Some models of quintessence predict an ever-accelerating expansion leading to a **"big rip" (p. 446)** that eventually would tear all the objects, even the atoms, of the universe apart.
- Corrected for acceleration, the observed value of the Hubble constant implies that the universe is 13.7 billion years old.
- The sudden inflation of the universe is hypothesized to have magnified tiny quantum mechanical fluctuations in the density of matter and energy. These very wide but very weak differences in density caused dark matter, followed by baryonic matter, to draw together and produce the present-day **large-scale structure (p. 446)** consisting of galaxy **superclusters (p. 446)** arranged in great walls and **filaments (p. 446)** outlining enormous **voids (p. 446).**
- Statistical observations of the large-scale structure of the universe and the size of irregularities in the cosmic microwave background confirm that it is flat and contains 4.6 percent baryonic matter, 23.3 percent dark matter, and 72.1 percent dark energy.
- The mass equivalent of dark energy added to dark matter and baryonic matter makes the observed density of the universe equal to the critical density, thereby confirming the prediction made by inflation theory that the universe is flat.

Review Questions

1. How does the darkness of the night sky tell you something important about the universe?
2. How can Earth be located at the center of the observable universe if you accept the cosmological principle?
3. Why can't an open universe have a center? How can a closed universe not have a center?
4. What evidence shows that the universe is expanding? What evidence shows that it began with a big bang?
5. Why couldn't atomic nuclei exist when the age of the universe was less than 2 minutes?
6. Why are measurements of the present density of the universe important?
7. How does the inflationary universe theory resolve the flatness problem? How does it resolve the horizon problem?
8. If the Hubble constant were really 100 instead of 70 km/s/Mpc, much of what astronomers understand about the evolution of stars and star clusters would have to be wrong. Explain why. (*Hint:* What would the age of the universe be?)
9. What is the evidence that the universe was very uniform during its first 400,000 years?
10. What is the difference between hot dark matter and cold dark matter? What difference does that make to cosmology?
11. What evidence can you cite that the expansion of the universe is accelerating?
12. What evidence can you cite that the universe is flat?
13. **How Do We Know?** Reasoning by analogy often helps make complicated systems or abstract ideas easier to understand. Why do you have to be careful when using analogies?
14. **How Do We Know?** The word *believe* has a meaning in normal conversation that does not really apply to scientific work. Why might cosmologists hesitate to use the word *believe* when they talk about the big bang?

Discussion Questions

1. Do you think Copernicus would have accepted the cosmological principle? Why or why not?
2. If you reject any model of the universe that has an edge in space because you can't comprehend such a thing, shouldn't you also reject any model of the universe that has a beginning? Isn't a beginning like an "edge" in time, or is there a difference?

Problems

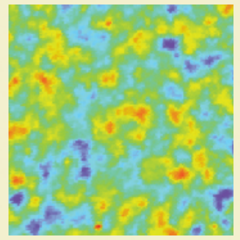
1. Use the data in Figure 19-4 to plot a velocity-distance diagram, calculate the Hubble constant *H*, and estimate the Hubble time.
2. If a galaxy is 9.0 Mpc away from Earth and recedes at 510 km/s, what is *H*? What is the Hubble time? How old would the universe be assuming space-time is flat and there has been no acceleration of the universe's expansion? How would acceleration change your answer?
3. What was the wavelength of maximum intensity for radiation from the gas of the big bang at the time of recombination? By what factor is that different from the wavelength of maximum intensity of the cosmic microwave background radiation observed now?
4. Pretend that galaxies are spaced evenly, 2 Mpc apart, and the average mass of a galaxy is 10^{11} solar masses. What is the average density of matter in the universe? (*Note:* The volume of a sphere is $\frac{4}{3}\pi r^3$, and the mass of the sun is 2×10^{30} kg.)
5. Figure 19-13 is based on an assumed Hubble constant of 70 km/s/Mpc. How would you change the diagram to fit a Hubble constant of 50 km/s/Mpc?
6. Hubble's first estimate of the Hubble constant was 530 km/s/Mpc. His distances were too small by a factor of about 7 because of a calibration error. If he had not had that calibration problem, what value for *H* would he have obtained?
7. What was the maximum age of the universe predicted by Hubble's first estimate of the Hubble constant (in Problem 6)?

Learning to Look

1. Explain why some of the galaxies in this photo have elongated, slightly curved images. What do such observations tell you about the universe?

NASA; ESA; L. Bradley (JHU); R. Bouwens (UCSC, Santa Cruz); H. Ford (JHU); and G. Illingworth (UCSC, Santa Cruz)

2. The image at the right shows irregularities in the background radiation. Why isn't the background radiation perfectly uniform? What does the size of these irregularities tell you?



Courtesy of the BOOMERANG Collaboration

Great Debates

1. *Fate of Humans?* The human race may face destruction by any of a number of events. Which will occur first, and which will occur last? Your choices are as follows: the sun's red giant stage; the return of Earth to protons, neutrons, and electrons; collision of the Milky Way Galaxy with Andromeda; asteroid or comet impact on Earth; global nuclear war; Yellowstone supervolcano explosion; the next ice age; runaway global warming; Earth's interior cooling and thus, elimination of Earth's magnetic field.
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Use information from your text and other sources about when these events could occur and/or their probability of occurring that supports your stand.
   c. Cite your sources.
2. *Dents in the CMB?* Recently, a claim has been made that the cosmic microwave background has a dent (or dents) in it due to other universe(s) colliding with ours. What color are the dents in the CMB? Is the claim about the collisions valid? Do we live in a multiverse (that is, a multiuniverse)?
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers in your debate.
   b. What's the evidence? Use information from your text and other sources including information on the small changes in the CMB pictures.
   c. Cite your sources.
3. *Olbers versus Digges.* Thomas Digges seems to have been the first to postulate Olbers's paradox. Should the name be changed to Digges's paradox? Is it really a paradox? What would you call it?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.
4. *The Cost of Fate.* Is spending taxpayer dollars to determine whether the universe is accelerating, nonaccelerating, or decelerating important? Would you vote for a candidate who wants to spend your tax dollars on the next space probe to determine the fate of the universe? If you were the candidate, which topic of astronomical research would you propose spending taxpayer dollars on?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Raisin Bread
- The First Three Minutes
- Decoupling of Radiation and Matter

An online, interactive way for you to explore astronomy, Virtual Astronomy Labs 2.0 allows you to have a hands-on lab experience from your computer. Enhance your understanding of the scientific method with the Virtual Astronomy Labs. Focusing on twenty of the most important concepts in astronomy, the labs offer you hands-on exercises that complement text topics.

Virtual Astronomy Lab 19: The Hubble Law

Among the many dizzying lessons in this chapter on cosmology, you learned that the universe has no center because it has no edge. The observable universe seems to have an edge—the limit of our vision with the most powerful telescopes—but that is a horizon, not an edge. Your horizon, whether viewed from a ship at sea or a car on the Great Plains, is the limit of your vision. But that horizon is not the limit of the world, which extends far beyond what you can see. And, of course, you perceive that you are at the exact center of your horizon. But so does everyone else—all observers find themselves at the exact center of their respective horizons.

The redshifts of galaxies are observed to be isotropic—the same in all directions as seen from Earth. Another way to say this is that you get the same Hubble constant, the same proportionality between galaxy distances and redshifts, when you observe in any direction. So are all the galaxies in the universe rushing away from our galaxy, and we happen to be at the center? Can you hear the spirit of Copernicus calling out to be careful, look for another explanation if you find yourself at what seems to be the center of the universe?

In this chapter, you learned to keep in mind that the universe has no center. This knowledge can help you avoid the common mistake of astronomy students (and some instructors!) of considering the redshifts of galaxies to be Doppler shifts. The Doppler effect occurs when a source and receiver of waves move through space relative to each other. Based on Einstein's well-tested general theory of relativity, scientists understand that distant galaxies are not really rushing away from us. Galaxies, including the Milky Way, have only small, random motions through space and stay at approximately the same "address" while the galaxies are being carried apart by the expansion of space-time.

As space-time stretches like a rubber sheet, it stretches all photons traveling through it. Photons from very distant galaxies take longer to reach us and get stretched more. That's why they have longer wavelengths and larger redshifts. Those redshifts are called "cosmological redshifts" to distinguish them from Doppler shifts. Astronomers living in any galaxy in the universe would see the same Hubble law and measure the same Hubble constant that expresses the rate of expansion of the universe.

Section 3 of Virtual Astronomy Lab 19, "The Hubble Law," allows you to study a model of the receding galaxies and reproduce Edwin Hubble's discovery of the expansion of the universe. You will also be able to show that it doesn't matter which galaxy you live in; the view, and the Hubble law, will be the same. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

Virtual Astronomy Lab 20: Fate of the Universe

You learned in this chapter that the phrase "expansion of the universe" means that space-time stretches and carries the galaxies away from each other. Astronomers can detect that expansion as redshifts in the spectra of galaxies, and they can measure the rate at which the universe is expanding that is expressed by the Hubble constant H. The Hubble constant is often written as H_0. (pronounced *H-nought*). That subscript "0" reminds you that we are referring to its present value. Astronomers have evidence that the universe is expanding faster now than in the "middle" past but slower than in the very distant past. H_0 represents the present rate of expansion that is constant across our local region of space-time.

If the universe actually had been expanding at a constant rate, you can easily estimate the age of the universe as $1/H_0$. Instead, suppose the universe's expansion is gradually speeding up. If that is so, you can't be as sure of its actual age or how it will expand in the future. This lab gives you a chance to see how the expansion of the universe depends on some basic assumptions.

The expansion of the universe depends on the curvature of space-time, and that curvature is determined by the average density of the universe. That density is made up of the density of matter, of course, but it also includes the density equivalent of all of the radiation, because radiation is energy and energy has a mass equivalent expressed by Einstein's famous equation $E = mc^2$. A third component of the universe's density is related to Einstein's cosmological constant. If Einstein was right to add the constant to his equations, then the empty vacuum of space contains some energy, called the vacuum energy. That vacuum energy also can be represented as a density. The overall density of the universe is the sum of those three densities: matter, energy, and vacuum energy.

Virtual Astronomy Lab 20, "Fate of the Universe," allows you to make some assumptions about the universe's three component densities, combine them with your estimate of the present value of the Hubble constant H_0, and see how the universe has expanded to the present day. That tells you the true age of the universe given your assumptions. Your calculations also predict the future fate of the expansion. You will see that the currently values measured for the three densities and H_0 lead to the conclusions that we live in a universe with open curvature and that the universe will expand forever. Sign in at http://login.cengagebrain.com to explore Virtual Astronomy Laboratories 2.0.

20

Astrobiology: Life on Other Worlds

Guidepost

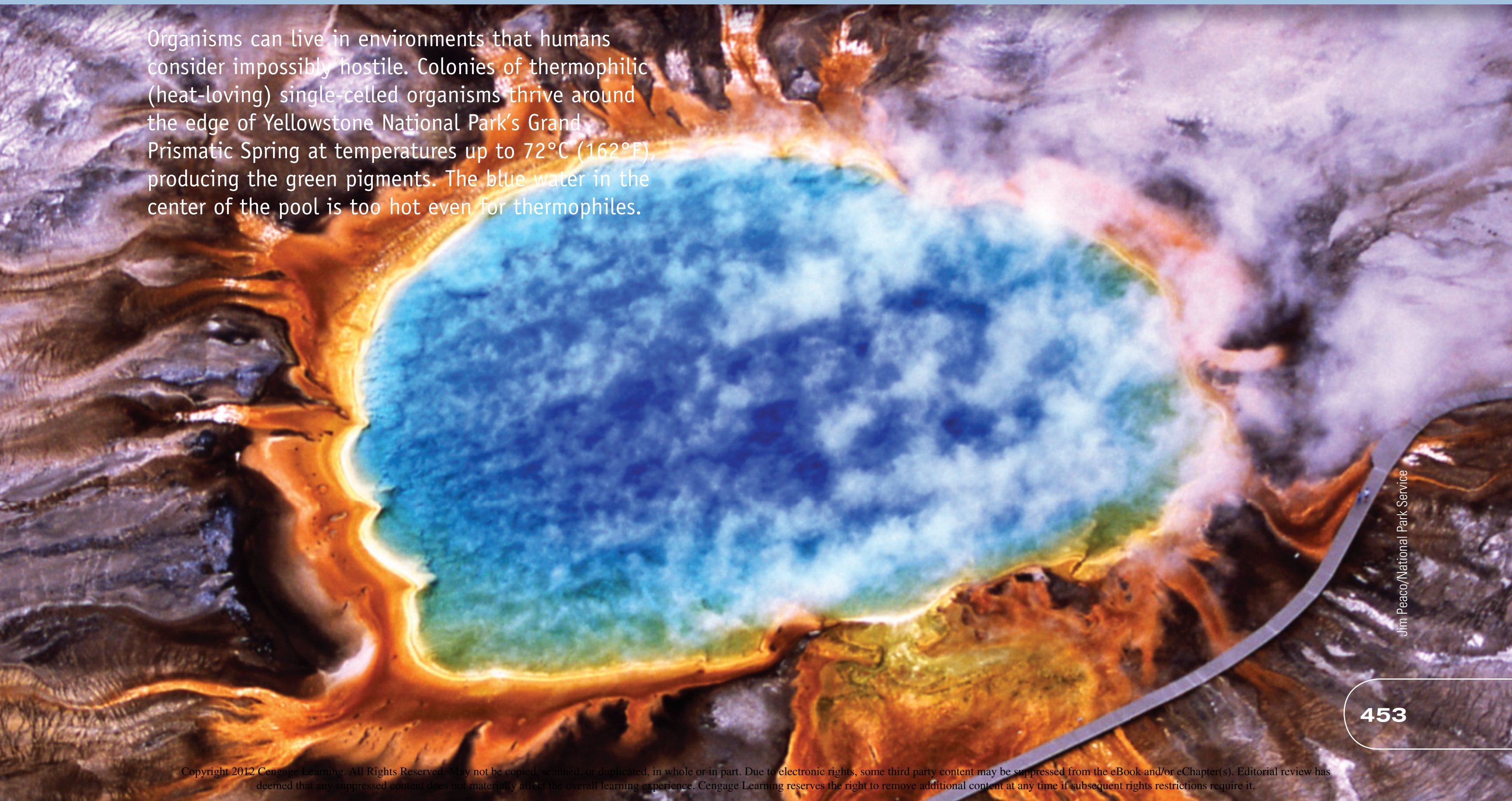
This chapter is either unnecessary or vital. If you believe that astronomy is the study of the physical universe above the clouds, then you are done; the previous 19 chapters completed your study of astronomy. But, if you believe that astronomy is the study not only of the physical universe but also of your role as a living being in the evolution of the universe, then everything you have learned so far from this book has been preparation for this final chapter.

As you read this chapter, you will encounter four important questions:

- **What is life?**
- **How did life originate on Earth?**
- **Does life exist on other worlds?**
- **Can humans communicate with intelligent beings on other worlds?**

You won't get more than the beginnings of answers to those questions here, but often in science asking a question is more important than getting an immediate answer.

You have explored the universe from the phases of the moon to the big bang, from the origin of Earth to the death of the sun. Astronomy is meaningful, not just because it is about the universe but because it is also about you. Now that you know astronomy, you see yourself and your world in a different way. Astronomy has changed you.



Organisms can live in environments that humans consider impossibly hostile. Colonies of thermophilic (heat-loving) single-celled organisms thrive around the edge of Yellowstone National Park's Grand Prismatic Spring at temperatures up to 72°C (162°F), producing the green pigments. The blue water in the center of the pool is too hot even for thermophiles.

Jim Peaco/National Park Service

Did I solicit thee from darkness
to promote me?

ADAM, TO GOD, IN JOHN MILTON'S *PARADISE LOST*

As a living thing, you have been promoted from darkness. The atoms of carbon, oxygen, and other heavy elements that are necessary components of your body did not exist at the beginning of the universe but were built up by successive generations of stars.

The elements from which you are made are common everywhere in the observable universe, so it is possible that life began on other worlds. Future explorers may find alien species different from any life on Earth. And it is possible some of those species have evolved to become intelligent. If so, perhaps those other civilizations will be detected from Earth.

Your goal in this chapter is to try to understand truly intriguing puzzles—the origin and evolution of life on Earth and what that tells you about whether there is life on other worlds (**How Do We Know? 20-1**). This new hybrid field of study is called **astrobiology.**

20-1 The Nature of Life

What is life? Philosophers have struggled with that question for thousands of years, and it is not possible to answer it completely in one chapter or even one book. An attempt at a general definition of what living things do, distinguishing them from nonliving things, might be: Life is a process by which an organism extracts energy from the surroundings, maintains itself, and modifies the surroundings to foster its own survival and reproduction.

One important observation is that all living things on Earth, no matter how apparently different, share certain characteristics in how they perform the process of life.

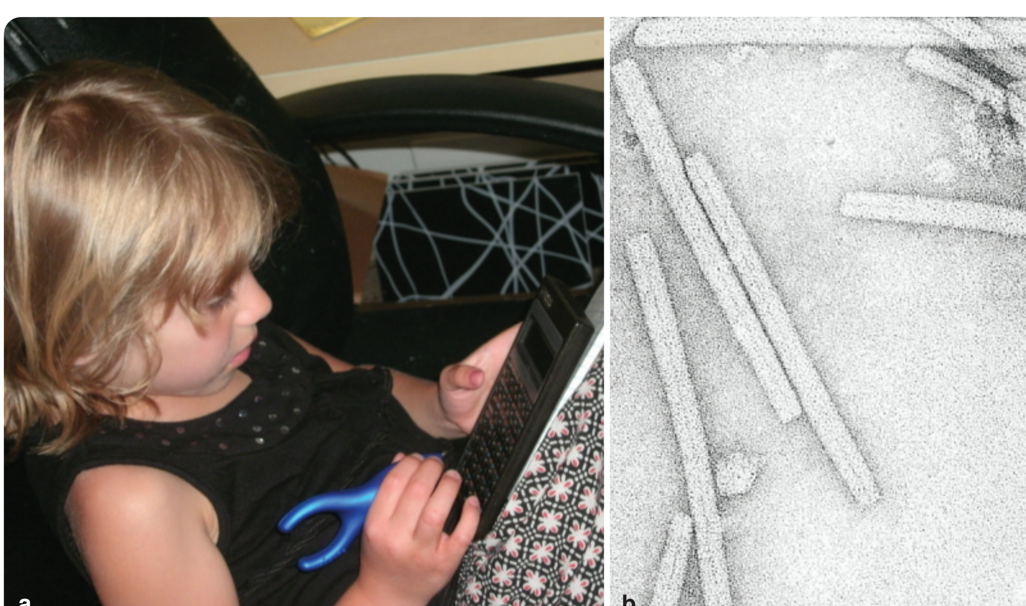
The Physical Bases of Life

The physical bases of life on Earth are carbon and water (■ Figure 20-1). Because of the way carbon atoms bond to each other and to other atoms, they can join into long, complex, stable chains that are capable, among many other feats, of storing and transmitting information. A large amount of information is necessary to control the activities and maintain the forms of living things. And, in all living things on Earth, the chemical reactions making, breaking, and combining carbon chains take place in liquid water within the cells of living organisms.

Is it possible that life on other worlds could be based on silicon instead of carbon? Silicon is right below carbon in the periodic table (Appendix Table A-16), which means that it shares many of carbon's chemical properties. But life based on silicon rather than carbon seems unlikely because silicon chains are harder to assemble and disassemble than their carbon counterparts and can't be as lengthy.

Is it possible that the chemistry of life on other worlds could take place in a setting other than water? Some alternatives such as methyl alcohol have been proposed. But carbon compounds dissolve especially easily in water. Also, water has exceptional properties such as a high heat capacity (resistance to temperature change) relative to other cosmically abundant substances that are liquid at the temperatures of planetary surfaces. It's not just because Earth scientists are themselves made of carbon and water that they think carbon and water are crucially important for the existence of life.

Science fiction has proposed even stranger life-forms based on, for example, electromagnetic



Jamie Backman

L. D. Simon/National Institute of Health

■ **Figure 20-1**

All living things on Earth are based on carbon chemistry in water. Even the long molecules that carry genetic information, DNA and RNA, have a framework defined by chains of carbon atoms. (a) Ana, a complex mammal, contains about 200 AU of DNA. (b) Each rodlike tobacco mosaic virus contains a single spiral strand of RNA about 0.01 mm long as its genetic material.

The Nature of Scientific Explanation

Must science and religion be in conflict? Science is a way of understanding the world around you, and at the heart of that understanding are explanations that science gives for natural phenomena. Whether you call these explanations stories, histories, hypotheses, or theories, they are attempts to describe how nature works based on evidence and intellectual honesty. While you may take these explanations as factual truth, you can understand that they are not the only explanations that describe the universe.

A separate class of explanations involves religion. For example, the Old Testament description of creation does not fit well with scientific observations, but it is a way of understanding the universe nonetheless. Religious explanations are based partly on faith rather than on strict rules of logic and evidence, and it is wrong to demand that they follow the same rules as scientific explanations. In the same way, it is wrong to demand that scientific explanations take into account religious beliefs. The so-called conflict between science and religion arises when people fail to recognize that science and religion are different ways of knowing about the universe.

Scientific explanations are quite compelling because science has been so successful at producing technological innovations that have changed the world you live in. From new vaccines, to digital music players, to telescopes that can observe the most distant galaxies, the products of the scientific process are all around you. Scientific explanations have provided tremendous insights into the workings of nature. Many people are attracted to the suggestion, made by evolutionary biologist Stephen Jay Gould and others, that religious explanations and scientific explanations should be considered as "separate magisteria" In other words, religion and science are devoted to different realms of the mystery of existence.

Science and religion offer differing ways of explaining the universe, but the two ways follow separate rules and cannot be judged by each other's standards. The trial of Galileo can be understood as a conflict between these two ways of knowing.

Fair use

Galileo's telescope gave him a new way to know about the universe.

fields and ionized gas, and none of these possibilities can be ruled out. Those hypothetical life-forms make for fascinating speculation, but for now they can't be studied systematically in the way that life on Earth can. This chapter is concerned with the origin and evolution of life as it is on Earth, based on carbon and water, not because of lack of imagination but because it is the only form of life about which we know anything.

Even carbon- and water-based life has its mysteries. What makes a lump of carbon-based molecules a living thing? An important part of the answer lies in the transmission of information from one molecule to another.

Information Storage and Duplication

Actions performed by living cells are carried out by molecules that are built within the cells. Cells must store recipes for making all those molecules as well as how and when to use them, and then somehow pass the recipes on to their offspring.

Study **DNA: The Code of Life** on pages 456–457 and notice three important points and seven new terms:

1. The chemical recipes of life are stored in each cell as information on *DNA (deoxyribonucleic acid)* molecules that resemble a ladder with rungs that are composed of chemical bases. The recipe information is expressed by the sequence of the rungs, providing instructions to guide specific chemical reactions within the cell.

2. The instructions stored in DNA are genetic information passed along to offspring. DNA instructions normally are expressed by being copied into a messenger molecule called *RNA (ribonucleic acid).* The RNA molecule travels to a location in the cell where its message causes a sequence of molecular units called *amino acids* to be connected into large molecules called *proteins*. Proteins serve as the cell's basic structural molecule or as *enzymes* that control chemical reactions.

3. The DNA molecule reproduces itself when a cell divides so that each new cell contains a copy of the original information. A sequence of DNA that composes one instruction is called a *gene.* Genes are organized into long coiled chains called *chromosomes*. The genes linked on one chromosome are normally passed on to offspring together.

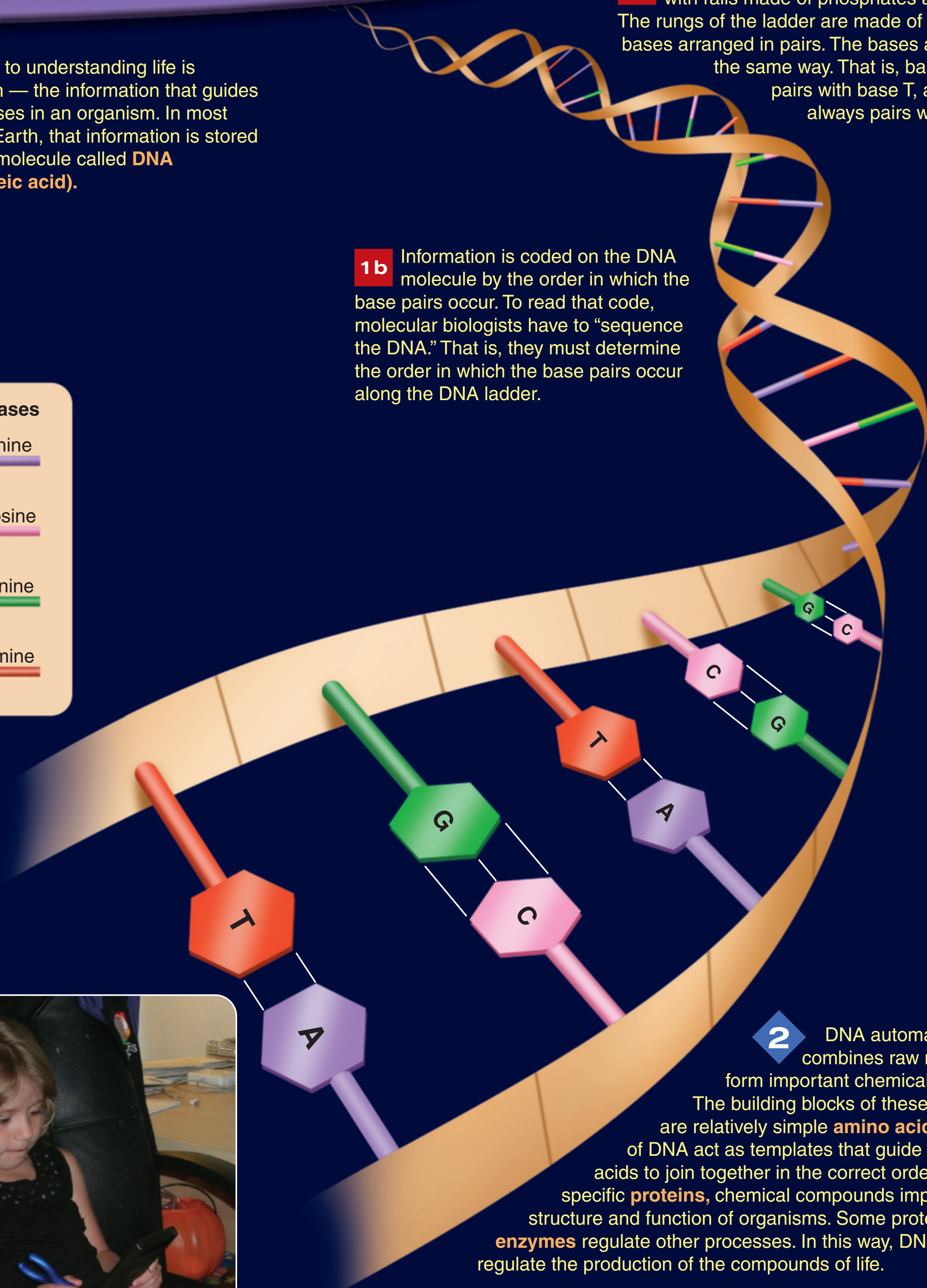
To produce viable offspring, a cell must be able to make copies of its DNA. Surprisingly, it is important for the continued existence of all life that the copying process includes mistakes.

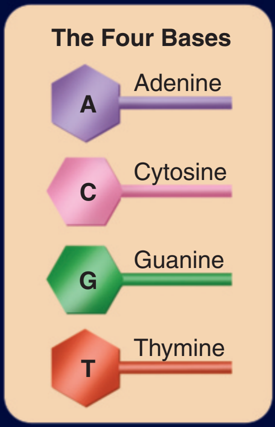
DNA: The Code of Life

1 The key to understanding life is information — the information that guides all of the processes in an organism. In most living things on Earth, that information is stored on a long spiral molecule called **DNA (deoxyribonucleic acid).**

1a The DNA molecule looks like a spiral ladder with rails made of phosphates and sugars. The rungs of the ladder are made of four chemical bases arranged in pairs. The bases always pair the same way. That is, base A always pairs with base T, and base G always pairs with base C.

1b Information is coded on the DNA molecule by the order in which the base pairs occur. To read that code, molecular biologists have to "sequence the DNA." That is, they must determine the order in which the base pairs occur along the DNA ladder.

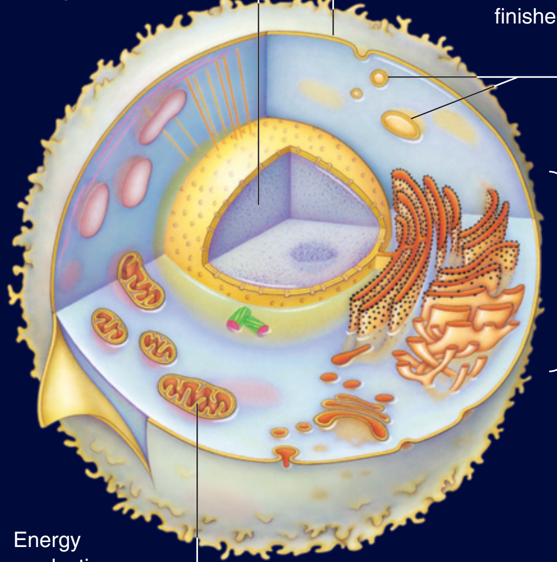




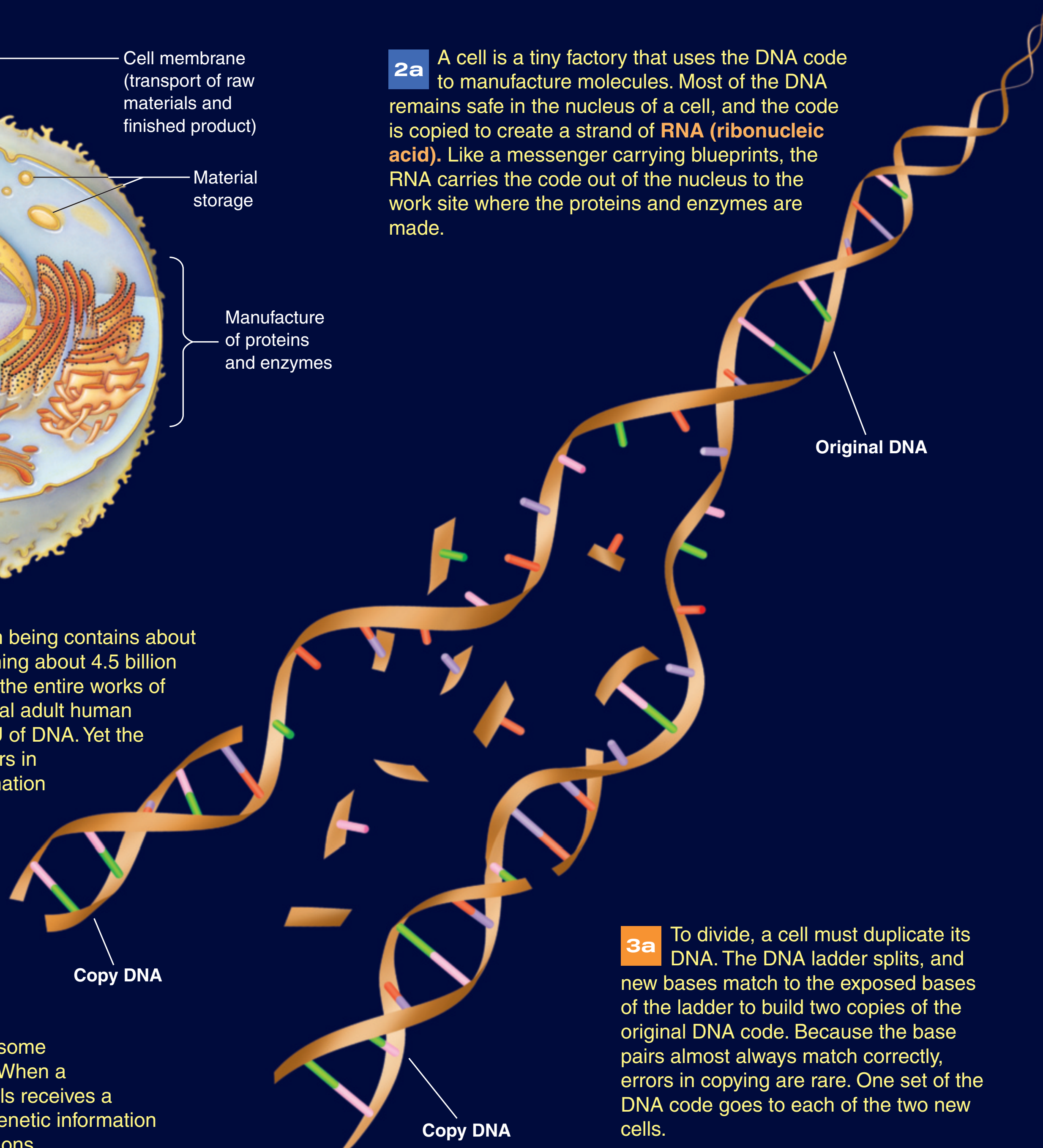
2 DNA automatically combines raw materials to form important chemical compounds. The building blocks of these compounds are relatively simple **amino acids.** Segments of DNA act as templates that guide the amino acids to join together in the correct order to build specific **proteins,** chemical compounds important to the structure and function of organisms. Some proteins called **enzymes** regulate other processes. In this way, DNA recipes regulate the production of the compounds of life.

Jamie Backman

The traits you inherit from your parents, the chemical processes that animate you, and the structure of your body are all encoded in your DNA. When people say "you have your mother's eyes," they are talking about DNA codes.



2a A cell is a tiny factory that uses the DNA code to manufacture molecules. Most of the DNA remains safe in the nucleus of a cell, and the code is copied to create a strand of **RNA (ribonucleic acid).** Like a messenger carrying blueprints, the RNA carries the code out of the nucleus to the work site where the proteins and enzymes are made.

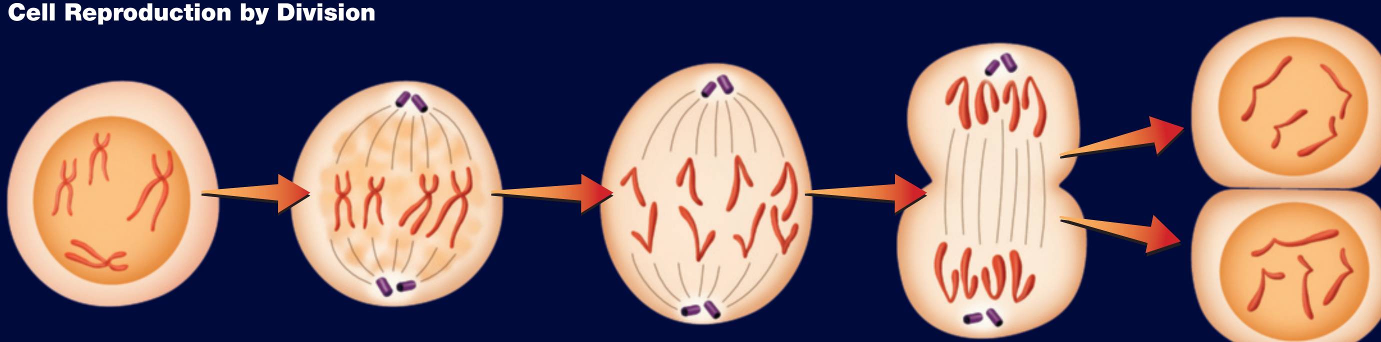


2b A single cell from a human being contains about 1.5 meters of DNA containing about 4.5 billion base pairs — enough to record the entire works of Shakespeare 200 times. A typical adult human contains a total of about 600 AU of DNA. Yet the DNA in each cell, only 1.5 meters in length, contains all of the information to create a new human. A clone is a new creature created from the DNA code found in a single cell.

3 DNA, coiled into a tight spiral, makes up the **chromosomes** that are the genetic material in a cell. A **gene** is a segment of a chromosome that controls a certain function. When a cell divides, each of the new cells receives a copy of the chromosomes, as genetic information is handed down to new generations.

3a To divide, a cell must duplicate its DNA. The DNA ladder splits, and new bases match to the exposed bases of the ladder to build two copies of the original DNA code. Because the base pairs almost always match correctly, errors in copying are rare. One set of the DNA code goes to each of the two new cells.

Cell Reproduction by Division



As a cell begins to divide, its DNA duplicates itself.

The duplicated chromosomes move to the middle.

The two sets of chromosomes separate, and . . .

the cell divides to produce . . .

two cells, each containing a full set of the DNA code.

Modifying the Information

Earth's environment changes continuously. To survive, species must change as their food supply, climate, or home terrain changes. If the information stored in DNA could not change, then life eventually would become extinct. The process by which life adjusts itself to changing environments is called **biological evolution.**

When an organism reproduces, its offspring receive a copy of its DNA. Sometimes external effects such as natural radiation alter the DNA during the parent organism's lifetime, and sometimes mistakes are made in the copying process, so that occasionally the copy is slightly different from the original. Offspring born with random alterations to their DNA are called **mutants.** Most mutations make no difference, but some mutations are fatal, killing the afflicted organisms before they can reproduce. In rare but vitally important cases, a mutation can actually help an organism survive.

These changes produce variation among the members of a species. All of the squirrels in the park may look the same, but they carry a range of genetic variation. Some may have slightly longer tails or faster-growing claws. These variations make almost no difference until the environment changes. For example, if the environment becomes colder, a squirrel with a heavier coat of fur will, on average, survive longer and produce more offspring than its normal contemporaries. Likewise, the offspring that inherit this beneficial variation will also live longer and have more offspring of their own. In contrast, squirrels containing DNA recipes for thin fur coats will gradually decrease in number. These differing rates of survival and reproduction are examples of **natural selection.** Over time, the beneficial variation becomes more common, and a species can evolve until the entire population shares the trait. In this way, natural selection adapts species to their changing environments by selecting, from the huge array of random variations, those that would most benefit the survival of the species.

It is a **Common Misconception** that evolution is random, but that is not true. The underlying mechanisms creating variation within each species may be random, but natural selection is not random because progressive changes in a species are directed by changes in the environment.

SCIENTIFIC ARGUMENT

Why is it important that errors occur in copying DNA?

Sometimes the most valuable scientific arguments are those that challenge what seems like common sense. It appears obvious that mistakes shouldn't be made in copying DNA, but in fact variation is necessary for long-term survival of a species. For example, the DNA in a starfish contains all the information the starfish needs to grow, develop, survive, and reproduce. The information must be passed on to the starfish's offspring for them to survive. That information must change, however, if the environment changes. A change in the ocean's temperature may kill the specific shellfish that the starfish eat. If none of the starfish are able to digest another kind of food—if all the starfish have exactly the same DNA—they all will die. But if a few starfish are born in each generation with the ability to make enzymes that can digest a different kind of shellfish, the species may be able to carry on.

Variations in DNA are caused both by external factors such as natural radiation and by occasional mistakes in the copying process. The survival of life depends on this delicate balance between mostly reliable reproduction and the introduction of small variations in DNA. Now build a new argument. **Why does the DNA copying process need to be mostly reliable?**

20-2 Life in the Universe

Life as we know it consists of just the single example of life on Earth. It is OK to think of all life on Earth as being just a single type of life because, as you learned in the previous section, all living things on Earth have the same physical basis: the same chemistry and the same genetic code alphabet. How life began on Earth and then developed and evolved into its present variety is the only solid information you have to work with when considering what might be possible on other worlds.

Everything currently known about life on Earth indicates that the same natural processes should lead to the origin of life on some fraction of other planets with liquid water. If there is life on other worlds, does it use DNA and RNA to carry the information for life processes, or different molecules playing the same role, or some radically different scheme? There is no way to know unless another example of life is found on another world. If and when that day comes, even if the non-Earthly life is a simple one-celled organism, the discovery will be one of the most important in the history of science. It will complete the journey of human understanding, begun in the Copernican revolution, of progressive realizations that Earth is not unique.

Origin of Life on Earth

It is obvious that the 4.5 billion chemical bases that make up human DNA did not come together in the right order just by chance. The key to understanding the origin of life lies in picturing the processes of evolution running "backwards." The complex interplay of environmental factors with the DNA of generation after generation of organisms drove some life-forms to become more sophisticated over time until they became the unique and specialized creatures alive today. Imagining this process in reverse leads to the idea that life on Earth began with extremely simple forms.

Biologists hypothesize that the first living things would have been carbon-chain molecules able to copy themselves. Of course, this is a scientific hypothesis for which you can seek evidence. What evidence exists regarding the origin of life on Earth?

Figure 20-2

(a) A fossil stromatolite from western Australia that is more than 3 billion years old, presenting some of the oldest evidence of life on Earth. Stromatolites are formed, layer upon layer, by mats of bacteria living in shallow water where they are covered repeatedly by sediments. (b) Artist's conception of a scene on the young Earth, 3 billion years ago, with stromatolite bacterial mats growing near the shores of an ocean.

Mural by Peter Sawyer

Chip Clark, National Museum of Natural History

The oldest fossils are the remains of sea creatures, and this indicates that life began in the sea. Identifying the oldest fossils is not easy, however. Ancient rocks from western Australia that are at least 3.4 billion years old contain features that biologists identify as **stromatolites,** fossilized remains of colonies of single-celled organisms (Figure 20-2). Fossils this old are difficult to recognize because the earliest living things did not contain easily preserved hard parts like bones or shells and because the individual organisms were microscopic. Thus the evidence, though scarce, indicates that simple organisms lived in Earth's oceans less than 1.2 billion years after Earth formed. Stromatolite colonies of microorganisms are more complex than individual cells, so you can imagine there probably were earlier, simpler organisms. Where did those first simple organisms come from?

An important experiment performed by Stanley Miller and Harold Urey in 1952 sought to recreate conditions on Earth before life began. The **Miller experiment** consisted of a sterile, sealed glass container holding water, hydrogen, ammonia, and methane, thought to simulate young Earth's atmosphere. An electric arc inside the apparatus made sparks to simulate the effects of lightning (Figure 20-3).

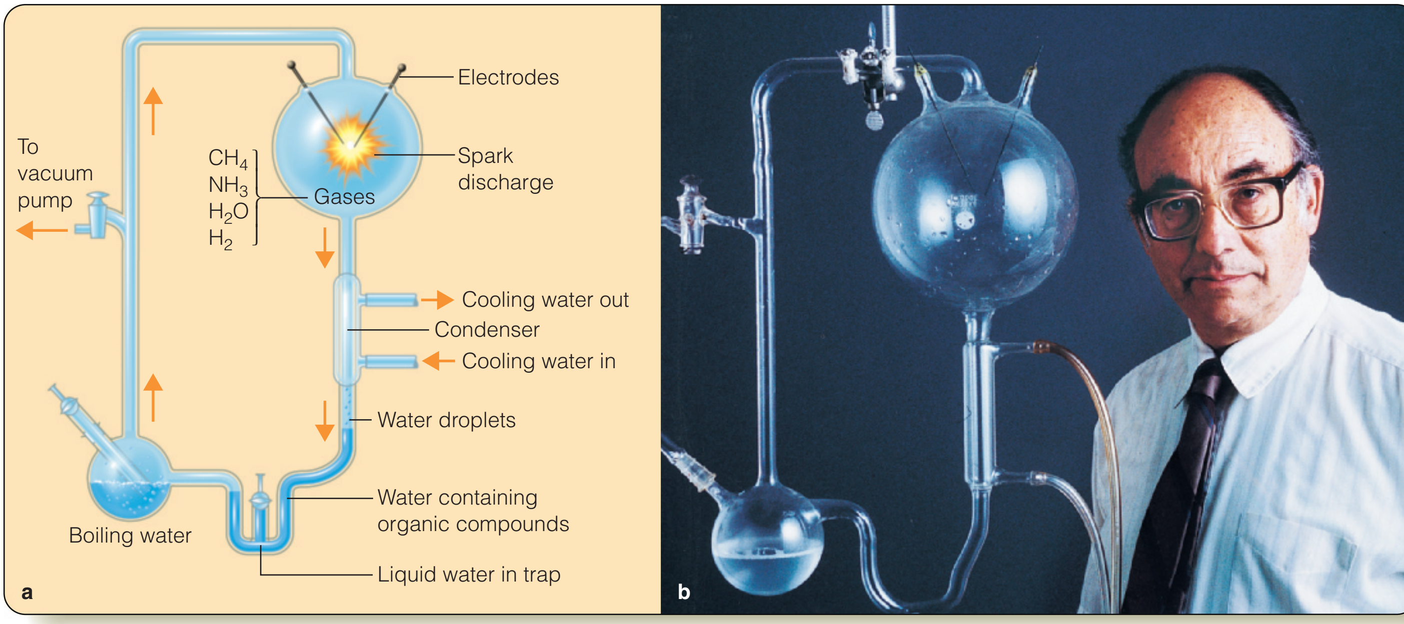
Miller and Urey let the experiment run for a week and then analyzed the material inside. They found that the interaction between the electric arc and the simulated atmosphere had produced many organic molecules from the raw material of the experiment, including such important building blocks of life as amino acids. (Recall that an organic molecule is simply a molecule with a carbon-chain structure and need not be derived from a living thing: "organic" does not necessarily imply "biological.")

When the experiment was run again using different energy sources such as hot silica to represent molten lava exposed to the atmosphere, similar molecules were produced. Even a source of ultraviolet radiation representing the small amount of UV in sunlight was sufficient to produce complex organic molecules.

Scientists are professionally skeptical about scientific findings (look back to "How Do We Know? 16-2"), and they have reevaluated the Miller-Urey experiment in light of new information. According to updated models of the formation of the solar system and Earth (see Chapters 8 and 9), Earth's early atmosphere probably consisted mostly of carbon dioxide, nitrogen, and water vapor instead of the mix of hydrogen, ammonia, methane, and water vapor assumed by Miller and Urey. When gases corresponding to the newer understanding of the early Earth atmosphere are processed in a Miller apparatus, lesser, but still significant, amounts of organic molecules are produced.

The Miller experiment is important because it shows that complex organic molecules form naturally in a wide variety of circumstances. Lightning, sunlight, and hot lava are just some of the energy sources that can naturally rearrange simple common molecules into the complex molecules that make life possible. If you could travel back in time, you would expect to find Earth's early oceans filled with a rich mixture of organic compounds called the **primordial soup.**

Many of these organic compounds would have been able to link up to form larger molecules. Amino acids, for example, can link together to form proteins by joining ends and releasing a water molecule (Figure 20-4). That reaction, however, does not proceed easily in a water solution. Scientists hypothesize that this step may have been more likely to happen on shorelines or in sun-warmed tidal pools where organic molecules from the primordial soup could have been concentrated by water evaporation. The production of large organic molecules may have been aided in such semidry environments by clay crystals acting as templates to hold the organic subunits close together.



© Cengage Learning 2014

Stanley Miller

■ Figure 20-3

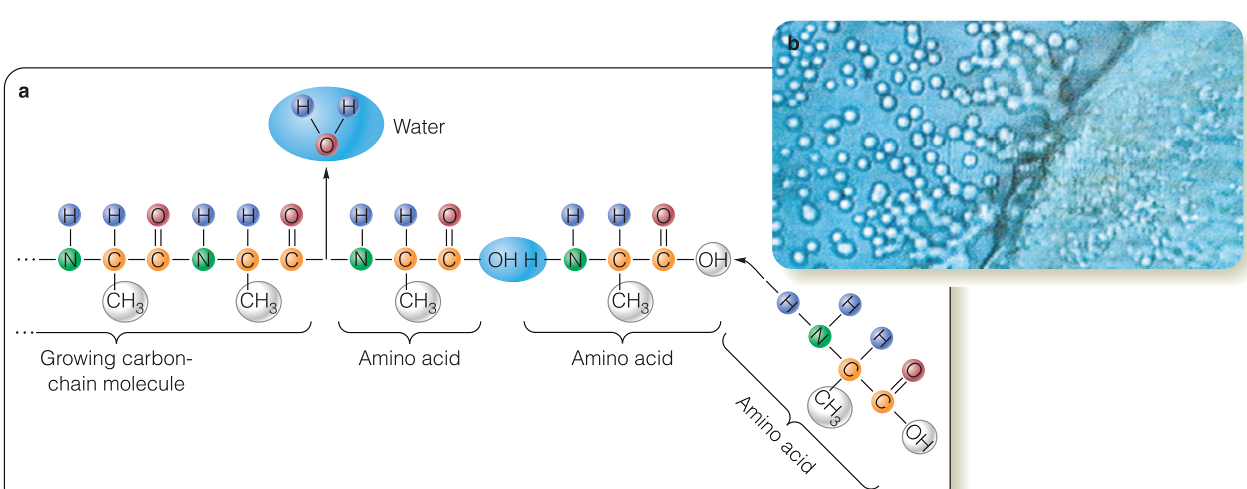
(a) The Miller experiment circulated gases through water in the presence of an electric arc. This simulation of primitive conditions on Earth produced many complex organic molecules, including amino acids, the building blocks of proteins. (b) Stanley Miller with a Miller apparatus.

These complex organic molecules were still not living things. Even though some proteins may have contained hundreds of amino acids, they did not reproduce, but rather linked and broke apart at random. Because some molecules are more stable than others, and some bond together more easily than others, scientists hypothesize that a process of **chemical evolution** eventually concentrated the various smaller molecules into the most stable larger forms. Eventually, according to the hypothesis, somewhere in the oceans, after sufficient time, a molecule formed that could copy itself. At that point, the natural selection and chemical evolution of molecules became the biological evolution of living things.

An alternate theory for the origin of life holds that reproducing molecules may have arrived here from space. Radio astronomers have found a wide variety of organic molecules in the interstellar medium, and similar compounds have been found inside meteorites (■ Figure 20-5). The Miller experiment showed how easy it is for complex organic molecules to form naturally from simpler

■ Figure 20-4

(a) Amino acids can link together via the release of a water molecule to form long carbon-chain protein molecules. The amino acid in this hypothetical example is alanine, one of the simplest. (b) When proteins cool in water, they often form microspheres, tiny globules with double-layered boundaries similar to cell membranes. Microspheres may have been an intermediate stage in the evolution of life between complex but nonliving molecules and living cells holding molecules reproducing genetic information.



© Cengage Learning 2014

Sydney Fox and Randall Grubbs

Chip Clark, National Museum of Natural History

Figure 20-5

A piece of the Murchison meteorite, a carbonaceous chondrite (see Chapter 12) that fell near Murchison, Australia, in 1969. Analysis of the interior of the meteorite revealed the presence of amino acids. Whether the first chemical building blocks of life on Earth originated in space is a matter of debate, but the amino acids found in meteorites illustrate how commonly amino acids and other complex organic molecules occur in the universe, even in the absence of living things.

compounds, so it is not surprising to find them in space. Although speculation is fun, the hypothesis that life arrived on Earth from space is presently more difficult to test than the hypothesis that Earth's life originated on Earth.

Whether the first reproducing molecules formed here on Earth or in space, the important thing is that they could have formed by natural processes. Scientists know enough about those processes to feel confident about them, even though some of the steps remain unknown.

The details of the origin of the first cells are unknown. The structure of cells may have arisen automatically because of the way molecules interact during chemical evolution. If a dry mixture of amino acids is heated, the acids form long, proteinlike molecules that, when poured into water, collect to form microscopic spheres that behave in ways similar to cells (pictured in the image on the opening page to this chapter). They have a thin membrane surface, they absorb material from their surroundings, they grow in size, and they divide and bud just as cells do. However, they contain no large molecule that copies itself, so they are not alive. The first reproducing molecule to be surrounded by a protective membrane, resulting in the first cell, would have gained an important survival advantage over other reproducing molecules.

Geologic Time and the Evolution of Life

Biologists infer that the first cells must have been simple single-celled organisms similar to modern bacteria. As you learned earlier, these kinds of cells are preserved in stromatolites, mineral formations produced by layers of bacteria and shallow ocean sediments (look back to Figure 20-2). Stromatolite fossils are found in rocks with radioactive ages of 3.4 billion years, and living stromatolites still form in some places today.

Stromatolites and other photosynthetic organisms would have begun adding oxygen, a product of photosynthesis, to Earth's early atmosphere. An oxygen abundance of only 0.1 percent would have created an ozone screen, protecting organisms from the sun's ultraviolet radiation and later allowing life to colonize the land.

Over the course of eons, the natural processes of evolution gave rise to stunningly complex **multicellular** life-forms with their own widely differing ways of life. It is a **Common Misconception** to imagine that life is too complex to have evolved from such simple beginnings. It is possible because small variations can accumulate, although that accumulation requires great amounts of time.

There is little evidence of anything more than simple organisms on Earth until about 540 million years ago, almost 3 billion years after the earliest signs of life, at which time fossils indicate that life suddenly developed into a wide variety of complex forms such as the trilobites (■ Figure 20-6). This sudden increase in complexity is known as the **Cambrian explosion** and marks the beginning of the Cambrian period.

If you represented the entire history of Earth on a scale diagram, the Cambrian explosion would be near the top of the column, as shown at the left of ■ Figure 20-7. The emergence of most animals familiar to you today, including fishes, amphibians, reptiles, birds, and mammals, would be crammed into the topmost part of the chart, above the Cambrian explosion.

If you magnify that portion of the diagram, as shown on the right side of Figure 20-7, you can get a better idea of when these events occurred in the history of life. Humanoid creatures have walked on Earth for about 4 million years. This is a long time by the standard of a human lifetime, but it makes only a narrow red line at the top of the diagram. All of recorded history would be a microscopically thin line at the very top of the column.

To understand just how thin that line is, imagine that the entire 4.6-billion-year history of the Earth has been compressed onto a yearlong video and that you began watching this video on January 1. You would not see any signs of life until March or early April, and the slow evolution of the first simple forms would take the next six or seven months. Suddenly, in mid-November, you would see the trilobites and other complex organisms of the Cambrian explosion.

Franklin & Marshall College/Grundy Observatory/Mike Seeds

Smithsonian and D. W. Miller

Figure 20-6

(a) Trilobites made their first appearance in the Cambrian oceans. The smallest were almost microscopic, and the largest were bigger than dinner plates. This example, about the size of your hand, lived 400 million years ago in an ocean floor that is now a limestone deposit in Pennsylvania. (b) In this artist's conception of a Cambrian sea bottom, Anomalocaris (rear at right center and looming at upper right), about human-hand-sized, had specialized organs including eyes, coordinated fins, gripping mandibles, and a powerful, toothed maw. Notice Opabinia at center right with its long snout.

You would see no life of any kind on land until November 28, but once life appeared it would diversify quickly, and by December 12 you would see dinosaurs walking the continents. By the day after Christmas they would be gone, and mammals and birds would be on the rise.

If you watched closely, you might see the first humanoid forms by late afternoon on New Year's Eve, and by late evening you could see humans making the first stone tools. The Stone Age would last until 11:59 PM, after which the first towns, and then cities, would appear. Suddenly things would begin to happen at lighting speed. Babylon would flourish, the pyramids would rise, and Troy would fall. The Christian era would begin 14 seconds before the New Year. Rome would fall; The Middle Ages and the Renaissance would flicker past. The American and French revolutions would occur one-and-a-half seconds before the end of the video.

By imagining the history of Earth as a yearlong video, you have gained some perspective on the rise of life. Tremendous amounts of time were needed for the first simple living things to evolve in the oceans. As life became more complex, new forms arose more and more quickly as the hardest problems—how to reproduce, how to take energy efficiently from the environment, how to move around—were "solved" by the process of biological evolution. The easier problems, like what to eat, where to live, and how to raise young, were managed in different ways by different organisms, leading to the diversity that is seen today.

Intelligence—that property that appears to set humans apart from other animals—may be a unique solution to an evolutionary problem posed to humanity's ancient ancestors. A smart animal is better able to escape predators, catch its prey, and feed and shelter itself and its offspring, so under certain conditions evolution is likely to naturally select for intelligence.

Life in Our Solar System

Could there be carbon-based life elsewhere in our solar system? As you learned earlier, liquid water seems to be a requirement of carbon-based life, necessary both as the medium for vital chemical reactions and to transport nutrients and wastes. It is not surprising that life developed in Earth's oceans and stayed there for billions of years before it was able to colonize the land.

It is difficult to pin down a range of environments and be sure that life based on carbon and water cannot exist outside those conditions, so long as there is even occasionally some liquid water present. Life has been found on Earth in places previously judged inhospitable, such as the bottoms of ice-covered lakes in Antarctica, far underground inside solid rock, and among the cinders at the summits of extinct volcanoes. An organism that can survive and even thrive in an extreme environment is called an **extremophile.** Scientists searching for life on other worlds must keep in mind Earth's extremophiles and the harsh conditions in which they thrive.

Many worlds in the solar system can be eliminated immediately as hosts for water-based life because liquid water is not possible there. The moon and Mercury are airless, and water would boil away into space immediately. Venus has traces of water vapor in its atmosphere, but it is too hot for liquid water to survive on the surface. The Jovian planets have deep atmospheres, and at a certain altitude it is likely that water condenses into liquid droplets. However, it seems unlikely that life could have originated there. The Jovian planets do not have solid surfaces (look back to Chapter 11), so isolated water droplets cannot mingle to mimic the original oceans of Earth, where organic molecules grew and interacted. Additionally, powerful downdraft

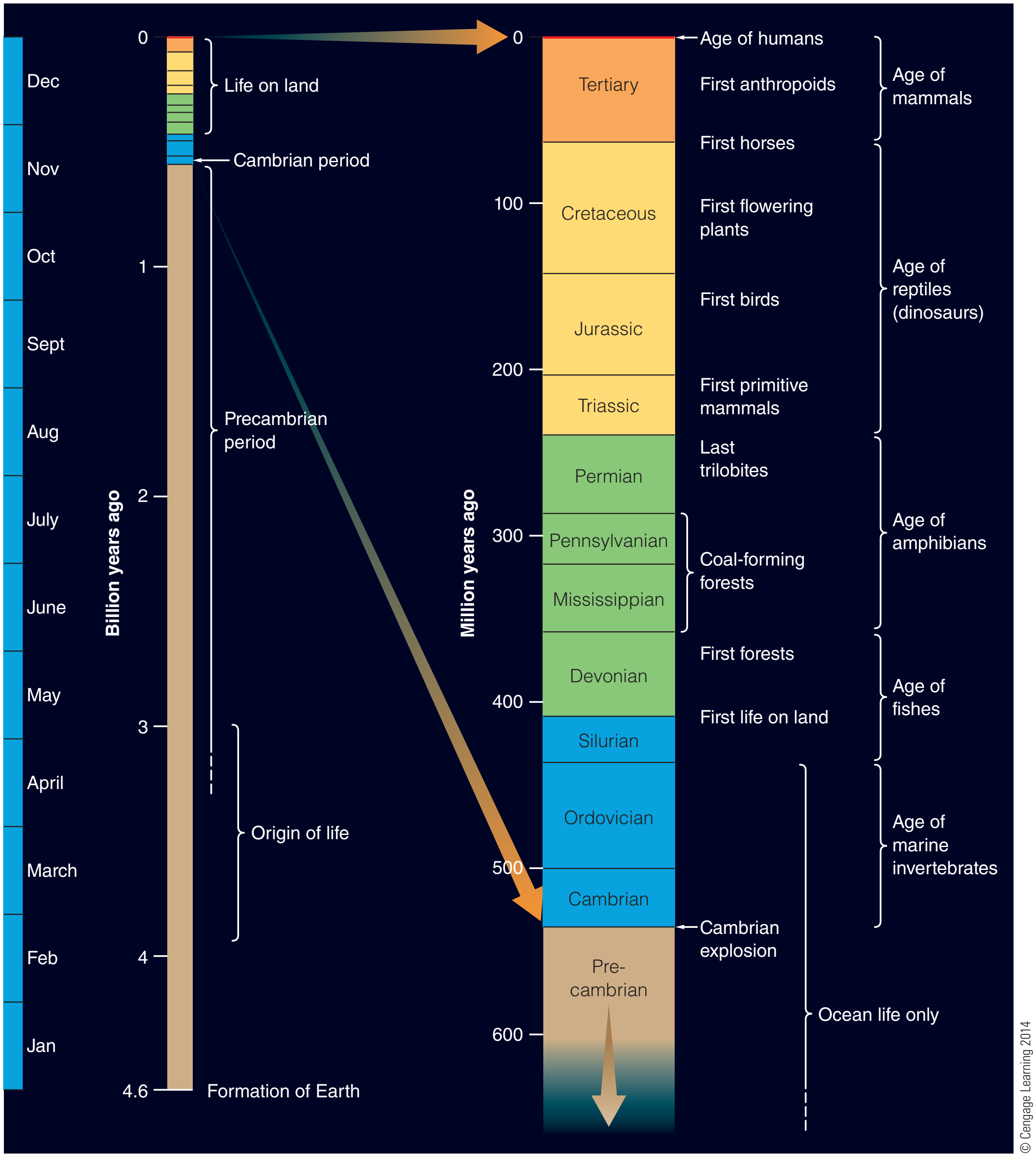


Figure 20-7

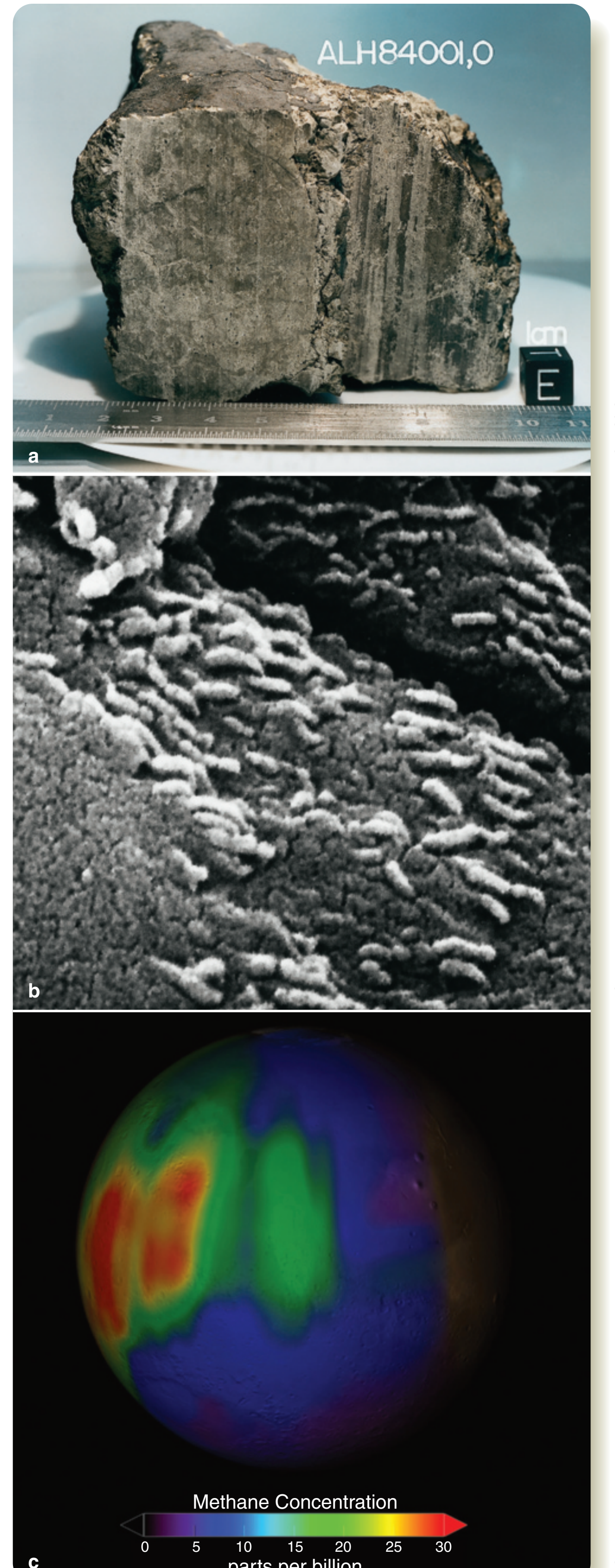
Complex life has developed on Earth only recently. If the entire history of Earth were represented in a time line (left), you have to examine the end of the line closely to see details such as life leaving the oceans and dinosaurs appearing. The age of humans would still be only a thin line at the top of your diagram. If the history of Earth were a yearlong videotape, humans would not appear until the last hours of December 31.

currents in the atmospheres of the giant planets would quickly carry any reproducing molecules that did form there into inhospitably hot lower regions.

As you learned in Chapter 11, at least one of the Jovian satellites could potentially support life. Jupiter's moon Europa appears to have a liquid-water ocean below its icy crust, and minerals dissolved in the water could provide a source of raw material for chemical evolution. Europa's ocean is kept liquid now by tidal heating. There are evidently also liquid zones inside Ganymede, Callisto, and Saturn's moon Enceladus. That can change as these moons interact gravitationally with their moons and their orbits change. All of them may have been frozen

■ **Figure 20-8**

(a) Meteorite ALH 84001 is one of a dozen meteorites known to have originated on Mars. Its name means this meteorite was the first one found in 1984 near Antarctica's Allan Hills. (b) A research group studying ALH 84001 claimed that the meteorite contains chemical and physical traces of ancient life on Mars, including what appear to be fossils of microscopic organisms. That evidence has not been confirmed, and the claim continues to be tested and debated. (c) A map of Mars with colors representing methane concentration in the atmosphere measured by spectrographs on Earth-based telescopes. Methane is most abundant in locations apart from volcanic regions, indicating the methane may be produced biologically.



solid at other times in their histories, which would probably have destroyed any living organism that had developed there.

Saturn's moon Titan is rich in organic molecules. You learned in Chapter 11 that sunlight converts the methane in Titan's atmosphere into organic smog particles that settle to the surface. The chemistry of life that could have evolved from those molecules and survived in Titan's lakes of methane is unknown. It is fascinating to consider possibilities, but Titan's extremely low temperature of −180°C (−290°F) would make chemical reactions so slow that life processes seem unlikely.

Water containing organic molecules has been observed venting from the south polar region of Saturn's moon Enceladus (look back again to Chapter 11). It is possible that life could exist in that water under Enceladus's crust, but the moon is quite small, and its tidal heating might operate only occasionally. Enceladus may not have had plentiful liquid water for the extended time necessary for the rise of life.

Aside from Earth, Mars is the most likely place for life to exist in the solar system because, as you learned in Chapter 10, there is a great deal of evidence that liquid water once flowed on its surface. Even so, results from searches for signs of life on Mars are not encouraging. The robotic spacecraft *Viking 1* and *Viking 2* landed on Mars in 1976 and tested soil samples for living organisms. Some of the tests had puzzling semipositive results that scientists hypothesize were caused by nonbiological chemical reactions in the soil. No evidence clearly indicates the presence of life or even of organic molecules in the Martian soil.

Life may have required special circumstances to start on Earth, but once it started, biological evolution allowed life to spread across Earth and adapt to a wide range of conditions. Eventually all niches—even extreme environments—became occupied. Scientists generally think this means that, if life begins on a planet, even if the entire environment of the planet later becomes inhospitable, as seems to have happened on Mars, some life could continue to survive. If life still exists on Mars, it may be hidden below ground where there may be liquid water and where UV radiation from the sun cannot penetrate.

There was a splash of news stories in the 1990s regarding supposed chemical and physical traces of life on Mars discovered inside a Martian meteorite found in Antarctica (■ Figure 20-8).

Scientists were excited by the announcement, but they employed professional skepticism and immediately began testing the evidence. Their results suggest that the unusual chemical signatures in the rock may have formed by processes that did not involve life. Tiny features in the rock that were originally thought to be fossils of ancient Martian microorganisms could possibly be nonbiological mineral formations instead.

In 2009 astronomers observing the Martian atmosphere announced detection of faint traces of methane, a substance made abundantly by living things on Earth. Methane would be destroyed in the Martian environment by solar UV radiation and chemical reactions, so the methane that is present must have been produced recently. There are also geological processes that can emit methane, but those processes are connected with volcanism. The map of Martian atmospheric methane concentrations in Figure 20-8c shows methane primarily in locations that are not volcanic provinces. It seems possible that organisms living in the Martian soil are making methane right now. This evidence regarding potential life on Mars remains highly controversial. Conclusive evidence of life on Mars may have to wait until a geologist from Earth can scramble down dry Martian streambeds and crack open rocks looking for fossils, or drill into the soil seeking signs of metabolizing microorganisms.

There is presently no compelling evidence for the existence of life in the solar system other than on Earth. Now your search will take you to distant planetary systems.

Life in Other Planetary Systems

Could life exist in other planetary systems? You already know that there are many different kinds of stars and that many of these stars have planetary systems. As a first step toward answering this question, you can try to identify the kinds of stars that seem most likely to have stable planetary systems where life could evolve.

If a planet is to be a suitable home for living things, it must be in a stable orbit around its sun. That is easy in a planetary system like our own, but planet orbits in binary star systems would be unstable unless the component stars are very close together or very far apart. Astronomers can calculate that, in binary systems with stars separated by intermediate distances of a few AU, the planets should eventually be swallowed up by one of the stars or ejected from the system. Half the stars in the galaxy are members of binary systems, and many of them are unlikely to support life on planets.

Moreover, just because a star is single does not necessarily make it a good candidate for sustaining life. Earth required perhaps as much as 1 billion years to produce the first cells and 4.6 billion years for intelligence to emerge. Massive stars that shine for only a few million years do not meet this criterion. If the time for development of life on Earth is typical, then stars more massive and luminous than about spectral type F5 last too short a time for complex life to develop. Main-sequence stars of types G, K, and M are the best candidates in terms of stellar lifetime.

The temperature of a planet is also important, and that depends on the type of star it orbits and its distance from the star. Astronomers have defined a **habitable zone** around a star as a region within which planets that orbit there have temperatures permitting the existence of liquid water. The sun's habitable zone extends from near the orbit of Venus to the orbit of Mars, with Earth right in the middle. A low-luminosity star has a small and narrow habitable zone, whereas a high-luminosity star has a large and wide one.

Stable planets inside the habitable zones of long-lived stars are the places where life seems most likely, but, given the tenacity and resilience of Earth's life-forms, there might be other, seemingly inhospitable, places in the universe where life exists. You should also note that three of the environments considered as possible havens for life—Europa, Titan, and Enceladus—are in the outer solar system, far outside the sun's conventionally defined habitable zone.

SCIENTIFIC ARGUMENT

What evidence indicates that life is possible on other worlds?

A good scientific argument involves careful analysis of evidence. Fossils on Earth show that life originated in the oceans at least 3.4 billion years ago, and biologists have outlined likely chemical processes that, over long time intervals, could have changed simple organic compounds into reproducing molecules inside membranes, the first simple life-forms. Meager fossil evidence indicates that life developed slowly at first. The pace of evolution quickened about half a billion years ago, when life took on complex forms. Later, life emerged onto the land and continued evolving rapidly into diverse forms. Intelligence is a relatively recent development on Earth: It is only a few million years old.

If this evolutionary process occurred on Earth, it seems reasonable that it could have occurred on other worlds as well. Earth-like worlds could be plentiful in the universe. Life may begin and eventually evolve to intelligence on any world where conditions are right. Now make a related argument. **What are the conditions you should expect on other worlds that host life?**

20-3 Intelligent Life in the Universe

Could intelligent life arise on other worlds? To try to answer this question, you can estimate the chances of any type of life arising on other worlds, then assess the likelihood of that life developing intelligence. If other civilizations exist, it is possible humans eventually may be able to communicate with them. Nature puts restrictions on the pace of such conversations, but the main problem lies in the unknown life expectancy of civilizations.

How Do We Know? 20-2

UFOs and Space Aliens

Has Earth been visited by aliens? If you conclude that there is likely to be life on other worlds, then you might be tempted to use UFO sightings as evidence to test your hypothesis. Scientists don't do this for two reasons.

First, the reputation of UFO sightings and alien encounters does not inspire confidence that these data are reliable. Most people hear of such events in grocery store tabloids, daytime talk shows, or sensational "specials" on viewer-hungry cable networks. You should take note of the low reputation of the media that report UFOs and space aliens. Most of these reports, like the reports that Elvis is alive and well, are simply made up for the sake of sensation, and you cannot use them as reliable evidence.

Second, the few UFO sightings that are not made up do not survive careful examination. Most are mistakes and unintentional misinterpretations, committed by honest people, of natural events or human-made objects. It is important to realize that experts have studied these incidents over many decades and found none that are convincing to the professional scientific community. In short, despite false claims to the contrary on TV shows, there is no dependable evidence that Earth has ever been visited by aliens.

In a way, that's too bad. A confirmed visit by intelligent creatures from beyond our solar system would answer many questions. It would be exciting, enlightening, and, like any real adventure, a bit scary. Most scientists would love to be part of such a discovery. But, scientists must professionally pay attention to what is supported by evidence rather than what might be thrilling. There is not yet any direct evidence of life on other worlds.

Flying saucers from space are fun to think about, but there is no evidence that they are real.

Travel Between the Stars

The distances between stars are almost beyond comprehension. The fastest human device ever launched, the *New Horizons* probe currently on its way to Pluto and the Kuiper belt (look back to Chapter 12), will take about 90,000 years to travel the distance to the nearest star, 4 light-years. The obvious way to overcome these huge distances is with tremendously fast spaceships, but even the closest stars are many light-years away.

Nothing can exceed the speed of light, and accelerating a spaceship close to the speed of light takes huge amounts of energy. Even if you travel slower than light, your rocket would still require massive amounts of fuel. If you wanted to pilot a spaceship with a mass of 100 tons (about the size of a fancy yacht) to the nearest star, and you traveled at half the speed of light so as to arrive in eight years, the trip would require 400 times as much energy as the entire United States consumes in a year. Don't even think about how much fuel the starship *Enterprise* needs.

These limitations not only make it difficult for humans to leave the solar system, but they would also make it difficult for aliens to visit Earth. Reputable scientists have studied "unidentified flying objects" (UFOs) and have never found any evidence that Earth is being visited or has ever been visited by aliens (**How Do We Know? 20-2**). Humans are unlikely ever to meet aliens face-to-face. However, communication by electromagnetic signals across interstellar distances takes relatively little energy.

Radio Communication

Nature puts restrictions on travel through space, and it also restricts the possibility of communicating with distant civilizations by radio. One restriction is based on simple physics: Radio signals are electromagnetic waves and travel at the speed of light. Due to the distances between the stars, the speed of radio waves would severely limit humanity's ability to carry on normal conversations with distant civilizations. Decades could elapse between asking a question and getting an answer.

So, rather than try to begin a conversation, one group of astronomers decided in 1974 to broadcast a message of greeting toward the globular cluster M13, 26,000 light years away, using the Arecibo radio telescope (look back to Figure 5-15). When the signal arrives 26,000 years in the future, alien astronomers may be able to understand it because the message is **anticoded,** meaning that it is intended to be decoded easily by beings about whom we know nothing except that they build radio telescopes. The message is a string of 1679 pulses and gaps. Pulses represent 1s, and gaps represent 0s. The string can be arranged in two dimensions in only two possible ways: as 23 rows of 73 or as 73 rows of 23. The second arrangement forms a picture containing information about life on Earth (■ Figure 20-9).

What are the chances that a signal like the Arecibo message would be heard across interstellar distances? Surprisingly, a radio dish the size of the Arecibo telescope, located anywhere in the

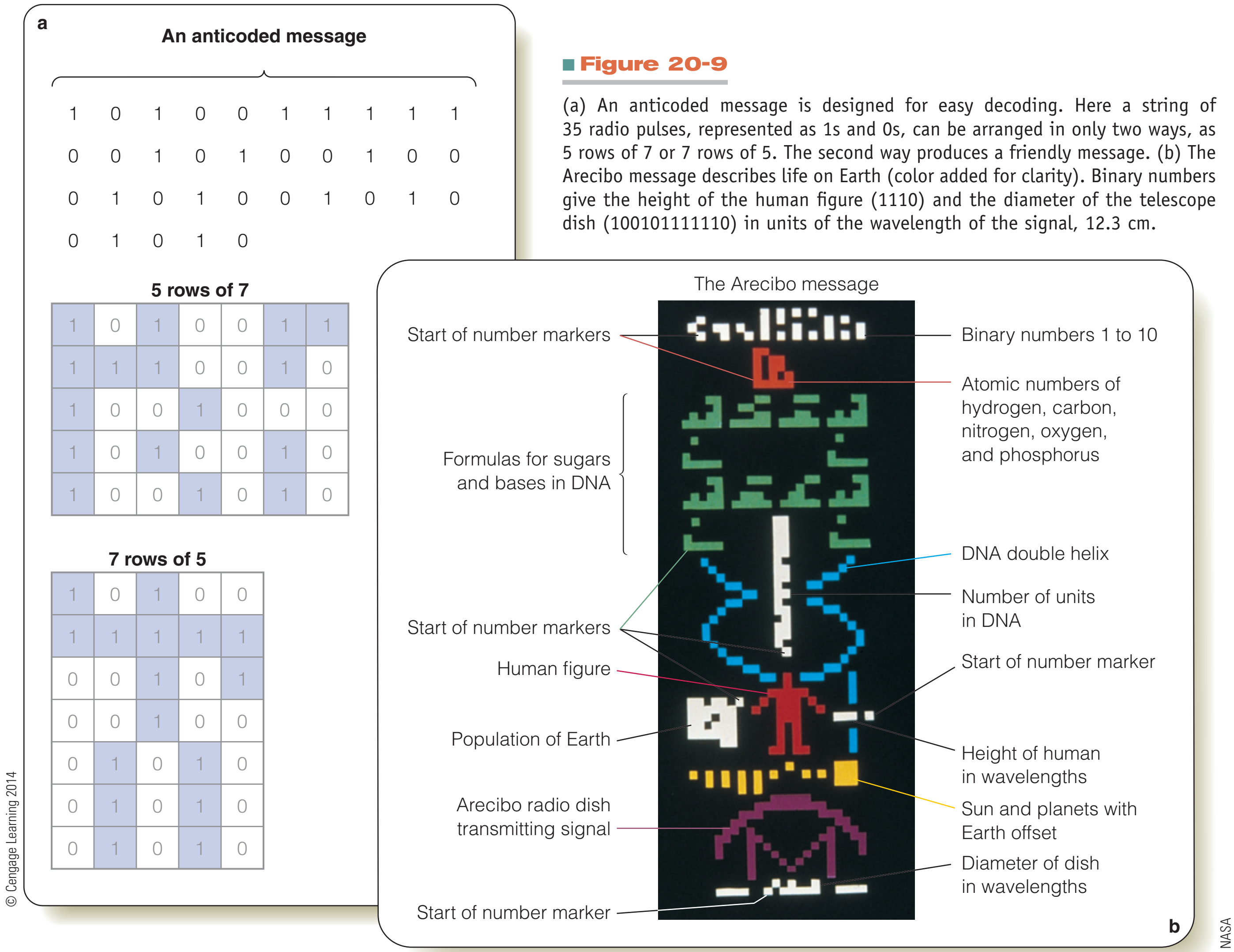


Figure 20-9

(a) An anticoded message is designed for easy decoding. Here a string of 35 radio pulses, represented as 1s and 0s, can be arranged in only two ways, as 5 rows of 7 or 7 rows of 5. The second way produces a friendly message. (b) The Arecibo message describes life on Earth (color added for clarity). Binary numbers give the height of the human figure (1110) and the diameter of the telescope dish (100101111110) in units of the wavelength of the signal, 12.3 cm.

Milky Way Galaxy, could detect the output from "our" Arecibo. The human race's modest technical capabilities already can put us into cosmic chat rooms.

Although the 1974 Arecibo beacon was the only powerful signal sent purposely from Earth to other star systems, Earth is sending out many other signals more or less accidentally. Short-wave radio signals, including TV and FM, have been leaking into space for the last 60 years or so. Any civilization within 60 light-years could already have detected Earth's civilization. That works both ways: Alien signals, whether intentional messages of friendship or the blather of their equivalent to daytime TV, could be arriving at Earth now. Groups of astronomers from several countries are pointing radio telescopes at the most likely stars and listening for alien civilizations.

Which channels should astronomers monitor? Signals with wavelengths longer than 30 cm would get lost in the background noise of our Milky Way Galaxy, while wavelengths shorter than about 1 cm are mostly absorbed in Earth's atmosphere. Between those wavelengths is a radio window that is open for communication. Even this restricted window contains millions of possible radio-frequency bands and is too wide to monitor easily, but astronomers may have thought of a way to narrow the search. Within this broad radio window lie the 21-cm spectral line of neutral hydrogen and the 18-cm line of OH (■ Figure 20-10). The interval between those lines has low background interference and is named the **water hole** because H plus OH yields water. Any civilizations sophisticated enough to do radio astronomy would know of these lines and might appreciate their significance in the same way as do Earthlings.

A number of searches for extraterrestrial radio signals have been made, and some are now under way. This field of study is known as **SETI, Search for Extra-Terrestrial Intelligence,** and it has generated heated debate among astronomers, philosophers, theologians, and politicians. You might imagine that the discovery of real alien intelligence would cause a huge change in humanity's worldview, akin to Galileo's discovery that the moons of Jupiter do not go around the Earth. Congress funded a NASA SETI search for a short time but ended support in the

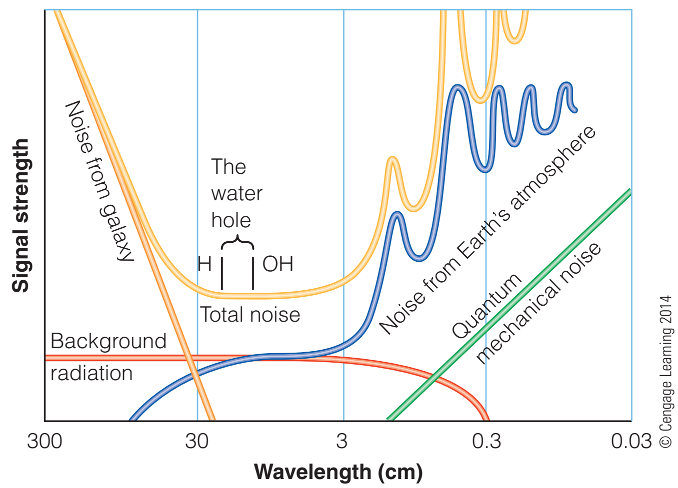


Figure 20-10

Radio noise from various astronomical sources and Earth's atmospheric opacity make it difficult to detect distant signals at wavelengths longer than 30 cm or shorter than 1 cm. In this range, wavelengths of radio emission lines from H atoms and from OH molecules mark a small wavelength range named the water hole that may be a likely channel for interstellar communication.

early 1990s. In fact, the annual cost of a major search is only about as much as a single Air Force helicopter, but much of the reluctance to fund searches stems from issues other than cost. Segments of the population, including some members of Congress, considered the idea of extraterrestrial beings as so outlandish that continued public funding for the search became impossible.

In spite of the controversy, the search continues. The NASA SETI project canceled by Congress was renamed Project Phoenix and completed using private funds. The SETI Institute, founded in 1984, managed Project Phoenix plus several other important searches and is currently building a new radio telescope array in northern California, in collaboration with the University of California, Berkeley, and partly funded by Paul Allen, one of the cofounders of Microsoft.

There is even a way for you to help with searches. The Berkeley SETI team (Note: they are separate from the SETI Institute), with the support of the Planetary Society, has recruited about 4 million owners of personal computers that are connected to the Internet. You can download a screen saver that searches data files from the Arecibo radio telescope for signals whenever you are not using the computer. For information, locate the seti//setiathome.ssl.berkeley.edu/.

The search continues, but radio astronomers struggle to hear anything against the worsening babble of radio noise from human civilization. Wider and wider sections of the electromagnetic spectrum are being used for Earthly communication, and this, combined with stray electromagnetic radiation from electronic devices including everything from computers to refrigerators, makes hearing faint radio signals difficult. It would be ironic if humans fail to detect faint radio signals from another world because our own world has become too noisy. One alternate search strategy is to look for rapid flashes of laser light at optical or near-infrared wavelengths. Such extraterrestrial signals, if they exist, would have the advantage of being easily distinguished from natural light sources but the disadvantage of being blocked by interstellar dust. Ultimately, the chance of success for any of the searches depends on the number of inhabited worlds in the galaxy.

How Many Inhabited Worlds?

Given enough time, the searches will find other worlds with civilizations, assuming that there are at least a few out there. If intelligence is common, scientists should find signals relatively soon—within the next few decades—but if intelligence is rare, it may take much longer.

Simple arithmetic can give you an estimate of the number of technological civilizations in the Milky Way Galaxy with which you might communicate, N_c. The formula proposed for discussions about N_c is named the **Drake equation** after the radio astronomer Frank Drake, a pioneer in the search for extraterrestrial intelligence. The version of the Drake equation presented here is modified slightly from its original form:

$$N_c = N_* \cdot f_P \cdot n_{HZ} \cdot f_L \cdot f_I \cdot f_S$$

N_* is the number of stars in our galaxy, and f_P represents the fraction of stars that have planets. The factor n_{HZ} is the average number of habitable planets in each planetary system—for the sake of the present discussion, the number of planets per system possessing substantial amounts of liquid water. The conventional habitable zone in our system contains Earth's orbit, and, arguably, also Venus and Mars. The *Kepler* mission has already revealed several Earth-size planets orbiting in the habitable zones around other stars. Moreover, Europa, Ganymede, Callisto, and Enceladus show that liquid water can exist due to tidal heating outside the conventional habitable zone. Thus, n_{HZ} may be larger than had been originally thought. The factor f_L is the fraction of suitable planets on which life begins, and f_I is the fraction of those planets where life evolves to intelligence.

The six factors on the righthand side of the Drake equation can be roughly estimated, with decreasing certainty as you proceed from left to right. The final factor is extremely uncertain. That factor f_S is the fraction of a star's life during which an intelligent species is communicative. If a society survives at a technological level for only 100 years, the chances of communicating with it are small. If most civilizations last only a short time, there may be none capable of transmitting during the cosmically short interval when Earthlings are capable of building radio telescopes to listen for them. On the other hand, a

Table 20-1 | The Number of Technological Civilizations per Galaxy

Estimates	Variables	Pessimistic	Optimistic
N_*	Number of stars per galaxy	2×10^{11}	2×10^{11}
f_P	Fraction of stars with planets	0.1	0.5
n_{HZ}	Number of planets per star that lie in habitable zone for longer than 4 billion years	0.01	1
f_L	Fraction of suitable planets on which life begins	0.01	1
f_I	Fraction of planets with life where life evolves to intelligence	0.01	1
f_S	Fraction of star's existence during which a technological society survives	10^{-8}	10^{-4}
N_c	Number of communicative civilizations per galaxy	2×10^{-4}	1×10^{7}

society that stabilizes and remains technologically capable for a long time is much more likely to be detected. For a star with a life span of 10 billion years, f_S might conceivably range from 10^{-8} for extremely short-lived societies to 10^{-4} for societies that survive for a million years. ■ Table 20-1 summarizes what many scientists consider a reasonable range of values for f_S and the other factors.

If the optimistic estimates are true, there could be a communicative civilization within a few tens of light-years from Earth. On the other hand, if the pessimistic estimates are true, Earth may be the only planet that is capable of communication within thousands of the nearest galaxies.

SCIENTIFIC ARGUMENT

Why does the number of civilizations that could be detected depend on how long civilizations survive at a technological level? This scientific argument depends on the timing of events. If you turned a radio telescope to the sky and scanned millions of frequency bands for many stars, you would be taking a snapshot of the universe at a particular time. Broadcasts from other civilizations must be arriving at that same time if they are to be detected. If most civilizations survive for a long time, there is a much greater chance that you will detect one of them in your snapshot than if civilizations tend to disappear quickly due to, for example, nuclear war or environmental collapse. There may be no civilizations in the galaxy interested in and capable of communicating with Earthlings during the cosmically short interval when we are also interested and capable.

The speed at which astronomers can search for signals is limited because computers must search many frequency intervals, but not all frequencies inside Earth's radio window are subject to intensive search. Now build a new argument to explain: **Why is the water hole an especially good frequency band in which to listen?**

What Are We? Matter and Spirit

There are over 4000 religions around the world, and nearly all hold that humans have a dual nature: We are physical objects made of atoms, but we are also spiritual beings. Science is unable to examine the spiritual side of existence, but it can tell us about our physical nature.

The matter you are made of appeared in the big bang and was cooked into a wide range of elements inside stars. Your atoms have been inside at least two or three generations of stars. Eventually, your atoms became part of a nebula that contracted to form our sun and the planets of the solar system.

Your atoms have been part of Earth for the last 4.6 billion years. They have been recycled many times through dinosaurs, stromatolites, fish, bacteria, grass, birds, worms, and other living things. You are using your atoms now, but when you are done with them, they will go back to Earth and be used again and again.

When the sun swells into a red giant star and dies in a few billion years, Earth's atmosphere and oceans will be driven away, and at least the outer few kilometers of Earth's crust will be vaporized and blown outward to become part of the nebula expanding around the white-dwarf remains of the sun. Your atoms are destined to return to the interstellar medium and will become part of future generations of stars and planets.

The message of astronomy is that humans are not just observers: We are participants in the universe. Among all of the galaxies, stars, planets, planetesimals, and bits of matter, humans are objects that can think, and that means we can understand what we are.

Is the human race the only thinking species? If so, we bear the sole responsibility to understand and admire the universe. The detection of signals from another civilization would demonstrate that we are not alone, and such communication would end the self-centered isolation of humanity and stimulate a reevaluation of the meaning of human existence. We may never realize our full potential as humans until we communicate with nonhuman intelligent life.

Summary

- **Astrobiology (p. 454)** is the scientific study of the likelihood that life exists elsewhere in the universe, and the study of life and its origin on Earth as it relates to life on other worlds.
- Life can be defined as a process that extracts energy from the surroundings, maintains an organism, and modifies the surroundings to promote the organism's survival.
- Living things have a physical basis—the arrangement of matter and energy that makes life possible. Life on Earth is based on carbon chemistry occurring in tiny bags of water (cells).
- Living things must have a controlling unit of information that can be passed to successive generations.
- Genetic information for life on Earth is stored in long carbon-chain molecules such as **DNA (deoxyribonucleic acid) (p. 456).**
- The DNA molecule stores information in the form of chemical bases linked together like the rungs of a ladder. Copied by the **RNA (ribonucleic acid) (p. 457)** molecule, the patterns of bases act as recipes for connecting together **amino acid (p. 456)** subunits to construct **proteins (p. 456),** including **enzymes (p. 456),** which are respectively the main structural and control components of the life process.
- The unit of heredity is a **gene (p. 457),** a piece or several pieces of DNA specifying construction of one or more protein molecules corresponding to one inheritable trait. Genes are connected in structures called **chromosomes (p. 457),** which are essentially single long DNA molecules.
- When a cell divides, the chromosomes split lengthwise and duplicate themselves so that each of the new cells can receive a copy of the genetic information.
- **Biological evolution (p. 458)** is the process by which life adjusts itself to changing environments.
- Errors in duplication or damage to the DNA molecule can produce **mutants (p. 458),** organisms that contain new DNA information and therefore have new properties. Variation in genetic codes can become widespread among individuals in a species. **Natural selection (p. 458)** determines which of these variations are best suited to survive, and the species evolves to fit its environment.
- Evolution is not random. Genetic variation is random, but natural selection is controlled by the environment.
- The oldest definitely identified fossils on Earth, structures called **stromatolites (p. 459)** that are composed of stacks of bacterial mats and sediment layers, are at least 3.4 billion years old. Those fossils provide evidence that life began in the oceans.
- Fossil evidence indicates that life began on Earth as simple single-celled organisms like bacteria and much later evolved into more complex, **multicellular (p. 461)** creatures.
- The **Miller experiment (p. 459)** shows that the chemical building blocks of life form naturally under a wide range of circumstances.
- Scientists hypothesize that **chemical evolution (p. 460)** occurred before biological evolution. Chemical evolution concentrated simple molecules into a diversity of larger stable organic molecules dissolved in the young Earth's oceans, but those molecules did not reproduce copies of themselves. The hypothetical organic-rich water is sometimes referred to as the **primordial soup (p. 459).** Biological evolution began when molecules developed the ability to make copies of themselves.
- Life-forms did not become large and complex until about 0.5 billion years ago, during what is called the **Cambrian explosion (p. 461).** Life emerged from the oceans only about 0.4 billion years ago, and human intelligence developed over the last 4 million (0.004 billion) years.
- Life as it is known on Earth requires liquid water and thus a specific range of temperatures. Organisms that thrive in extreme environments are called **extremophiles (p. 462).**
- No other planet in our solar system appears to harbor life at present. Most are too hot or too cold, although life might have begun on Mars before the planet became too cold and dry. If so, life conceivably could persist today on Mars in limited hospitable environments.
- Liquid water exists, and therefore Earth-like life is at least possible, under the surfaces of Jupiter's moons Europa and Ganymede and Saturn's moon Enceladus. Saturn's moon Titan has abundant organic compounds but does not have liquid water.
- Because the origin of life and its evolution into intelligent creatures took so long on Earth, scientists do not consider middle- and upper-main-sequence stars, which shine for astronomically short time spans, as likely homes for life.
- Main-sequence G and K stars and possibly some of the M stars are thought to be likely candidates to host planets with life.
- The **habitable zone (p. 465)** around a star, within which planets can have liquid water on their surfaces, may be larger than scientists had expected, given the wide variety of living things now found in extreme environments on Earth. Tidally heated moons orbiting large planets could have liquid water at any distance from a star.
- Because of distance, speed, and fuel, travel between the stars seems almost impossible for humans or for aliens who might visit Earth.
- Communication between planetary systems using electromagnetic signals such as radio waves or laser beams may be possible, but a real conversation would be difficult because of long travel times for such signals.
- Broadcasting a radio (or light) beacon of pulses would distinguish the signal from naturally occurring emission and identify the source as a technological civilization. The signal can be **anticoded (p. 466)** in the hope that another civilization can decode it.
- One good part of the radio spectrum for communication is called the **water hole (p. 467),** the wavelength range from the 21-cm spectral line of hydrogen to the 18-cm line of OH. Even so, millions of radio wavelengths need to be tested to fully survey the water hole for a given target star.
- Sophisticated searches are now under way to detect radio transmissions from civilizations on other worlds, but such **SETI (Search for Extra-Terrestrial Intelligence) (p. 467)** programs are hampered by limited computer power and radio noise pollution from human civilization.
- The number of civilizations in our galaxy that are at a technological level and able to communicate while humans are listening can be estimated by the **Drake equation (p. 468).** This number is limited primarily by the lifetimes of their and our civilizations.

Study and Review

Review Questions

1. If life is based on information, what is that information?
2. How does the DNA molecule produce a copy of itself?
3. What would happen to a life-form if the genetic information handed down to offspring was copied extremely inaccurately? How would that endanger the future of the life-form?
4. What would happen to a life-form if the information handed down to offspring was always the same? How would that endanger the future of the life-form?
5. Give an example of natural selection acting on new DNA patterns to select the most advantageous characteristics.
6. What evidence do scientists have that life on Earth began in the sea?
7. Why do scientists generally think that liquid water is necessary for the origin of life?
8. What is the difference between chemical evolution and biological evolution?
9. What is the significance of the Miller experiment?
10. How does intelligence make a creature more likely to survive?
11. Why are upper-main-sequence (high-luminosity) stars unlikely sites for intelligent civilizations?
12. Why is it reasonable to suspect that travel between stars is nearly impossible?
13. How does the stability of technological civilizations affect the probability that Earth can communicate with them?
14. What is the water hole, and why is it a good “place” to search for extraterrestrial civilizations?
15. Why is it difficult to anticode a message?
16. **How Do We Know?** How do science and religion have complementary explanations of the world?
17. **How Do We Know?** Why are scientists confident that Earth has never been visited by aliens?

Discussion Questions

1. Do you expect that hypothetical alien recipients of the Arecibo message will be able to decode it? Why or why not?
2. How do you think the detection of extraterrestrial intelligence would be received by the public? Would it be likelier to upset, or confirm, humans’ beliefs about themselves and the world?
3. What do you think it would mean if decades of careful searches for radio signals for extraterrestrial intelligence turn up nothing?

Problems

1. A single human cell encloses about 1.5 m of DNA, containing 4.5 billion base pairs. What is the spacing between these base pairs in nanometers? That is, how far apart are the rungs on the DNA ladder?
2. If you represent Earth’s history by a line 1 m long, how long a segment would represent the 400 million years since life moved onto the land? How long a segment would represent the 4-million-year history of human life?
3. If a human generation, the average time from birth to childbearing, has been 20 years long, how many generations have passed in the last 1 million years?
4. If a star must remain on the main sequence for at least 4 billion years for life to evolve to intelligence, what is the most massive a star can be and still possibly harbor intelligent life on one of its planets? (*Hints:* See Reasoning with Numbers 14-1 and Appendix Table A-7.)
5. If there are about 1.4×10^{-4} stars like the sun per cubic light-year, how many lie within 100 light-years of Earth? (*Hint:* The volume of a sphere is $\frac{4}{3}\pi r^3$.)
6. Mathematician Karl Gauss suggested planting forests and fields in gigantic geometric figures as signals to possible Martians that intelligent life exists on Earth. If Martians had telescopes that could resolve details no smaller than 1 arc second, how large would the smallest element of Gauss’s signal have to be for it to be visible at Mars’s closest approach to Earth? (*Hint:* See Appendix Table A-10, and use the small-angle formula, Chapter 3.)
7. If you detected radio signals with an average wavelength of 20 cm and suspected that they came from a civilization on a distant Earth-like planet, roughly how much of a change in wavelength should you expect to see because of the orbital motion of the distant planet?
8. Calculate the number of communicative civilizations per galaxy using your own estimates of the factors in Table 20-1.

Learning to Look

1. The star cluster shown in the image below contains cool red giants and main-sequence stars from hot blue stars all the way down to red dwarfs. Discuss the likelihood that planets orbiting any of these stars might be home to life. (*Hint:* Estimate the age of the cluster.)

NASA

2. If you could search for life in the galaxy shown in the image below, would you look among disk stars, or halo stars? Discuss the factors that influence your decision.

ESO

Enhanced Web Assign

New tutorials written exclusively for this text by the author may be assigned in Enhanced WebAssign:

- Earth Calendar
- The Drake Equation

Great Debates

1. *Top of the Food Chain, Beyond Humans?* Say a large asteroid or comet impact is about to eliminate over 99 percent of human and other surface life on Earth. What living creature would evolve to become the top of the food chain on land? Use the information in this chapter on where life originated on Earth and how humans originally became the top of the food chain on land.
   a. Use at least three vocabulary words from the text correctly in your debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

2. *The Moon and the Drake Equation.* Whether the moon is necessary for humans to have evolved on Earth is a question for debate. Earth is stabilized by its large moon, such that Earth has mild seasons and a long-term climate that is relatively stable. Is a large moon necessary in the Drake equation for a planet to become inhabited with intelligent life?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

3. *Grandma's Time Casket.* Sorry to say, your grandmother has died. She has requested the your parents preside over her will. Your grandmother was a lover of astrobiology and requested that when she died that she be buried in a time capsule along with various possessions so future generations millions of years from now can learn about life in the past. Grandma's time casket and funeral will cost $50,000, and she has given you and your parents $50,000 to see to her wishes. No legal document signed by your grandmother forces you to bury her and her possessions in this manner. Your father wants to give her a normal funeral and keep the rest of the money. Your mother wants to honor her wishes. Your parents want your advice. How do you counsel your parents?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

4. *Alien Life Exists!* You heard on the news that a NASA scientist found evidence of life on a meteorite. He claims that the fossilized remains of alien microorganisms are not so different than those found on Earth. The find, if true, could mean that life began as the solar system began, and life on Earth did not necessarily begin on Earth. Do you believe the find? Why or why not? You are one of 5000 people who have been asked to cite your beliefs on this find. What do you write?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

5. *UFO Conspiracy.* Do you believe governments are involved in a conspiracy to cover up the existence of UFOs? If so, which governments? Should governments report findings of UFOs to the general public or only to a select few? Why or why not? If you say a select few, who is included in the inner circle?
   a. Use at least three vocabulary words from the text correctly in your written debate, underline each, and cite the page and paragraph numbers.
   b. What's the evidence? Find additional sources to support your stand.
   c. Cite your sources.

6. *Favorite Subject.* What is the most interesting topic that you learned from this textbook? What will you never forget? Is this the most interesting subject in the book? Would your classmates agree? Did you pass any of the knowledge you gained in this course onto someone else not in the class?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. Embed a picture about your subject, and explain the picture.
   c. Cite your sources.

7. *Least Favorite Subject.* What is the least interesting topic that was covered in this textbook? What topic did you feel you will never use or see again on TV or in the news?
   a. Use at least three vocabulary words from the text correctly, underline each, and cite the page and paragraph numbers.
   b. Embed a picture about your subject, and explain the picture.
   c. Cite your sources.

Afterword

The aggregate of all our joys and sufferings, thousands of confident religions, ideologies and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilizations, every king and peasant, every young couple in love, every hopeful child, every mother and father, every inventor and explorer, every teacher of morals, every corrupt politician, every superstar, every supreme leader, every saint and sinner in the history of our species, lived there on a mote of dust, suspended in a sunbeam.

CARL SAGAN (1934–1996)

Earth photographed by *Voyager 1* from the edge of the solar system.

NASA/JPL

Our journey together is over, but before we part company, ponder one final time the primary theme of this book—humanity's place in the physical universe. Astronomy gives us some comprehension of the workings of stars, galaxies, and planets, but its greatest value lies in what it teaches us about ourselves. Now that you have surveyed astronomical knowledge, you can better understand your own position in nature.

To some, the word *nature* conjures up visions of furry rabbits hopping about in a forest glade. To others, nature is the blue-green ocean depths, and still others think of nature as windswept mountaintops. As diverse as these images are, they are all Earth-bound. Having studied astronomy, you can see nature as a beautiful mechanism composed of matter and energy, interacting according to simple rules, forming galaxies, stars, planets, mountaintops, ocean depths, forest glades, and people.

Perhaps the most important astronomical lesson is that humanity is a small but important part of the universe. Most of the universe is probably lifeless. The vast reaches between the galaxies appear to be empty of all but the thinnest gas, and stars are much too hot to preserve the chemical bonds that seem necessary for life to survive and develop. It seems that only on the surfaces of a few planets, where temperatures are moderate, can atoms link together in special ways to form living matter.

If life is special, then intelligence is precious. The universe must contain many planets devoid of life, planets where sunlight has shined unfelt for billions of years. There may also exist planets on which life has developed but has not become complex, planets where the wind stirs wide plains of grass and rustles through dark forests. On some planets, creatures resembling Earth's insects, fish, birds, and animals may watch the passing days only dimly aware of their own existence. It is intelligence, human or otherwise, that gives meaning to the landscape.

Science is the process by which Earth's intelligence has tried to understand the physical universe. Science is not the invention of new devices or processes. It does not create home computers, cure the mumps, or manufacture plastic spoons—those are engineering and technology, the adaptation of scientific understanding for practical purposes. Science is the understanding of nature, and astronomy is that understanding on the grandest scale. Astronomy is the science by which the universe, through its intelligent lumps of matter, tries to understand its own existence.

As the primary intelligent species on this planet, we are the custodians of a priceless gift—a planet filled with living things. This is especially true if life is rare in the universe. In fact, if Earth is the only inhabited planet, our responsibility is overwhelming. We are the only creatures who can take action to preserve the existence of life on Earth; ironically, our own actions are the most serious hazards.

The future of humanity is not secure. We are trapped on a tiny planet with limited resources and a population growing faster than our ability to produce food. We have already driven some creatures to extinction and now threaten others. We are changing the climate of our planet in ways we do not fully understand. Even if we reshape our civilization to preserve our world, the sun's evolution will eventually destroy Earth.

This may be a sad prospect, but a few factors are comforting. First, everything in the universe is temporary. Stars die, galaxies die; perhaps the entire universe will someday end. As part of a much larger whole, we are reminded that our distant future is limited. Second, we have a few billion years to prepare, and a billion years is a very long time. Only a few million years ago, our ancestors were starting to walk upright and communicate. A billion years ago, our ancestors were microscopic organisms living in the oceans. To suppose that a billion years hence there will be beings resembling today's humans, or that humans will still be the dominant intelligence on Earth, or that humans will even exist, are ultimately conceits.

Our responsibility is not to save our race for all eternity but to behave as dependable custodians of our planet, preserving it, admiring it, and trying to understand it. That calls for drastic changes in our behavior toward other living things and a revolution in our attitude toward our planet's resources. Whether we can change our ways is debatable—humanity is far from perfect in its understanding, abilities, or intentions. However, you must not imagine that we, and our civilization, are less than precious. We have the gift of intelligence, and that is the finest thing this planet has ever produced.

> We shall not cease from exploration
> And the end of all our exploring
> Will be to arrive where we started
> And know the place for the first time.
> *—T. S. Eliot, "Little Gidding"*

Appendix A
Units and Astronomical Data

Introduction

A system of units is based on the three fundamental units for length, mass, and time. By international agreement, there is a preferred subset of metric units known as the *Système International d'Unités* (SI units), commonly called the metric system, that is based on the meter, kilogram, and second. Other quantities, such as density and force, are derived from these fundamental units.

U.S. residents generally use the (British) Imperial system of units (officially used only in the United States, Liberia and Myanmar, but, ironically, not in Great Britain). For example, in Imperial units the fundamental unit of length is the foot, composed of 12 inches.

SI units employ the decimal system. For example, a meter is composed of 100 centimeters. Because the metric system is a decimal system, it is easy to express quantities in larger or smaller units as is convenient. You can give distances in centimeters, meters, kilometers, and so on. The prefixes specify the relation of the unit to the meter. Just as a cent is 1/100 of a dollar, so a centimeter is 1/100 of a meter. A kilometer is 1000 m, and a kilogram is 1000 g. The meanings of the commonly used prefixes are given in ■ Table A-1.

Fundamental and Derived SI Units

The three fundamental SI units define the rest of the units, as given in ■ Table A-2.

The SI unit of force is the newton (N), named after Isaac Newton. It is the force needed to accelerate a 1 kg mass by 1 m/s^2, or the force roughly equivalent to the weight of an apple at Earth's surface. The SI unit of energy is the joule (J), the energy produced by a force of 1 N acting through a distance of 1 m. A joule is roughly the energy in the impact of an apple falling off a table.

Exceptions

Units can help you in two ways. They make it possible to make calculations, and they can help you to conceive of certain quantities. For calculations, the metric system is far superior, and it is used for calculations throughout this book.

In SI units, density should be expressed as kilograms per cubic meter, but no human hand can enclose a cubic meter, so that unit does not help you grasp the significance of a given density. This book refers to density in grams per cubic centimeter. A gram is roughly the mass of a paperclip, and a cubic centimeter is the size of a small sugar cube, so you can easily conceive of a density of 1 g/cm^3, roughly the density of water. This is not a bothersome departure from SI units because you will not have to make complex calculations using density.

For conceptual purposes this book expresses some quantities in both SI and Imperial units. Instead of saying the average person would weigh 133 N on the moon, it might be more helpful to some readers for that weight to be expressed as 30 lb. In such cases the Imperial form is given in parentheses after the SI form. For example, the radius of the moon is 1738 km (1080 mi).

Conversions

To convert from one metric unit to another (from meters to kilometers, for example), you have only to look at the prefix. However, converting from metric to English or English to metric is more complicated. The conversion factors are given in ■ Table A-3.

■ **Table A-1 | Metric Prefixes**

Prefix	Symbol	Factor
Mega	M	10^6
Kilo	k	10^3
Centi	c	10^{-2}
Milli	m	10^{-3}
Micro	μ	10^{-6}
Nano	n	10^{-9}

■ **Table A-2 | SI *(Système International)* Metric Units**

Quantity	SI Unit
Length	Meter (m)
Mass	Kilogram (kg)
Time	Second (s)
Force	Newton (N)
Energy	Joule (J)

Table A-3 | Conversion Factors Between British and Metric Units

1 inch = 2.54 centimeters	1 centimeter = 0.394 inch
1 foot = 0.3048 meter	1 meter = 39.37 inches = 3.28 feet
1 mile = 1.609 kilometers	1 kilometer = 0.6214 mile
1 slug = 14.59 kilograms	1 kilogram = 0.06852 slug
1 pound = 4.448 newtons	1 Newton = 0.2248 pound
1 foot-pound = 1.356 joules	1 Joule = 0.7376 foot-pound
1 horsepower = 745.7 joules/s	1 Joule/s = 0.001341 horsepower
	1 Joule/s = 1 Watt

Example: The radius of the moon is 1738 km. What is this in miles? Table A-3 indicates that 1.000 mile equals 1.609 km, so

$$1738 \text{ km} \times \frac{(1.000 \text{ mi})}{(1.609 \text{ km})} = 1080 \text{ mi}$$

Temperature Scales

In astronomy, as in most other sciences, temperatures are expressed on the Kelvin scale, although the centigrade (or Celsius) scale is also used. The Fahrenheit scale commonly used in the United States is not used in scientific work.

The centigrade scale refers temperatures to the freezing point of water (0°C) and to the boiling point of water (100°C). One degree Centigrade is 1/100, the temperature difference between the freezing and boiling points of water; thus the prefix *centi.* The centigrade scale is also called the Celsius scale after its inventor, the Swedish astronomer Anders Celsius (1701–1744).

Temperatures on the Kelvin scale are measured in Celsius degrees from absolute zero (−273.15°C), the temperature of an object that contains no extractable heat. In practice, no object can be as cold as absolute zero, although laboratory apparatuses have reached temperatures lower than 10^{-6} K. The Kelvin scale is named after the Scottish mathematical physicist William Thomson, Lord Kelvin (1824–1907).

The Fahrenheit scale fixes the freezing point of water at 32°F and the boiling point at 212°F. Named after the German physicist Gabriel Daniel Fahrenheit (1686–1736), who made the first successful mercury thermometer in 1720, the Fahrenheit scale is used routinely only in the United States.

It is easy to convert temperatures from one scale to another using the information given in Table A-4.

Table A-4 | Temperature Scales and Conversion Formulas

	Kelvin (K)	Centigrade (°C)	Fahrenheit (°F)
Absolute zero	0 K	−273°C	−460°F
Freezing point of water	273 K	0°C	32°F
Boiling point of water	373 K	100°C	212°F

Conversions:

$K = °C + 273$

$°C = \frac{5}{9}(°F - 32)$

$°F = \frac{9}{5}(°C) + 32$

Powers of 10 Notation

Powers of 10 make writing very large numbers much simpler. For example, the nearest star is about 43,000,000,000,000 km from the sun. Writing this number as 4.3×10^{13} km is much easier.

Very small numbers can also be written with powers of 10. For example, the wavelength of visible light is about 0.0000005 m. In powers of 10 this becomes 5×10^{-7} m.

The powers of 10 used in this notation appear below. The exponent tells you how to move the decimal point. If the exponent is positive, move the decimal point to the right. If the exponent is negative, move the decimal point to the left. For example, 2.0000×10^3 equals 2000, and 2×10^{-3} equals 0.002.

$$\begin{aligned} 10^5 &= 100{,}000 \\ 10^4 &= 10{,}000 \\ 10^3 &= 1{,}000 \\ 10^2 &= 100 \\ 10^1 &= 10 \\ 10^0 &= 1 \\ 10^{-1} &= 0.1 \\ 10^{-2} &= 0.01 \\ 10^{-3} &= 0.001 \\ 10^{-4} &= 0.0001 \end{aligned}$$

If you use scientific notation in calculations, be sure you correctly enter the numbers into your calculator. Not all calculators accept scientific notation, but those that can have a key labeled EXP, EEX, or perhaps EE that allows you to enter the exponent of ten. To enter a number such as 3×10^8, press the keys 3 EXP 8. To enter a number with a negative exponent, you must use the change-sign key, usually labeled +/− or CHS. To enter the number 5.2×10^{-3}, press the keys 5.2 EXP +/− 3. Try a few examples.

To read a number in scientific notation from a calculator you must read the exponent separately. The number 3.1×10^{25} may appear in a calculator display as 3.1 25 or on some calculators as 3.1 10^{25}. Examine your calculator to determine how such numbers are displayed.

Astronomy Units and Constants

Astronomy, and science in general, is a way of learning about nature and understanding the universe. To test hypotheses about how nature works, scientists use observations of nature. The tables that follow contain some of the basic observations that support science's best understanding of the astronomical universe. Of course, these data are expressed in the form of numbers, not because science reduces all understanding to mere numbers, but because the struggle to understand nature is so demanding that science must use every valid means available. Quantitative thinking—reasoning mathematically—is one of the most powerful techniques ever invented by the human brain. Thus, these tables are not nature reduced to mere numbers but numbers supporting humanity's growing understanding of the natural world around us.

Table A-5 | Astronomical Constants

Velocity of light *(c)*	$= 3.00 \times 10^8$ m/s
Gravitational constant *(G)*	$= 6.67 \times 10^{-11}$ m^3/s^{2}kg
Mass of H atom	$= 1.67 \times 10^{-27}$ kg
Mass of Earth $(M_\oplus)$	$= 5.97 \times 10^{24}$ kg
Earth equatorial radius $(R_\oplus)$	$= 6.38 \times 10^3$ km
Mass of sun $(M_\odot)$	$= 1.99 \times 10^{30}$ kg
Radius of sun $(R_\odot)$	$= 6.96 \times 10^8$ m
Solar luminosity $(L_\odot)$	$= 3.83 \times 10^{26}$ j/s
Mass of moon	$= 7.35 \times 10^{22}$ kg
Radius of moon	$= 1.74 \times 10^3$ km

Table A-6 | Units Used in Astronomy

1 Angstrom (Å)	$= 10^{-8}$ cm
	$= 10^{-10}$ m
	$= 10$ nm
1 astronomical unit (AU)	$= 1.50 \times 10^{11}$ m
	$= 93.0 \times 10^6$ mi
1 light-year (ly)	$= 6.32 \times 10^4$ AU
	$= 9.46 \times 10^{15}$ m
	$= 5.88 \times 10^{12}$ mi
1 parsec (pc)	$= 2.06 \times 10^5$ AU
	$= 3.09 \times 10^{16}$ m
	$= 3.26$ ly
1 kiloparsec (kpc)	= 1000 pc
1 megaparsec (Mpc)	= 1,000,000 pc

Table A-7 | Properties of Main-Sequence Stars

Spectral Type	Absolute Visual Magnitude (M_v)	$L^{1,2}$	Temp. (K)	λ_{max} (nm)	Mass[1]	Radius[1]	Average Density (g/cm^3)
O5	−5.7	620,000	42,000	69	60	12	0.03
B0	−4.0	61,000	30,000	97	18	7.4	0.04
B5	−1.2	1,100	15,000	191	5.9	3.9	0.1
A0	0.7	73	9800	296	2.9	2.4	0.2
A5	2.0	18	8200	354	2.0	1.7	0.4
F0	2.7	8.8	7300	397	1.6	1.5	0.5
F5	3.5	4.6	6600	436	1.4	1.3	0.6
G0	4.4	2.1	5900	488	1.05	1.1	0.8
G2	4.7	1.0	5200	558	1.0	1.0	0.1
G5	5.1	0.7	5600	521	0.9	0.9	0.8
K0	5.9	0.6	5200	563	0.8	0.8	1.2
K5	7.4	0.3	4400	657	0.7	0.7	1.8
M0	8.8	0.1	3800	755	0.5	0.6	2.2
M5	12.3	0.01	3200	914	0.2	0.3	10

[1]Luminosity, mass, and radius are given in terms of the sun's luminosity, mass, and radius.

[2]Luminosity is computed from radius and temperature.

■ Table A-8 | The Brightest Stars[1,2]

Star	Name	Apparent Visual Magnitude (m_v)	Distance[3] (pc)	Absolute Visual Magnitude[4] (M_v)	Spectral Type
	(Sun)	−26.74		4.8	G2 V
α CMa	Sirius	−1.47	2.6	1.5	A1 V
α Car	Canopus	−0.72	95	−5.6	F0 II
α Boo	Arcturus	−0.04	11	−0.2	K1.5 III
α Cen A	Rigil Kentaurus	−0.01	1.3	4.4	G2 V
α Lyr	Vega	0.03	7.7	0.6	A0 V
β Ori	Rigel	0.12	260	−7.0	B8 Iab
α CMi	Procyon	0.34	3.5	2.6	F5 IV-V
α Ori	Betelgeuse	0.42	150	−5.5	M2 Iab
α Eri	Achernar	0.50	43	−2.7	B3 Vpe
β Cen	Hadar	0.60	120	−4.8	B1 III
α1 Aur	Capella A	0.71	13	0.1	G8 III
α Aql	Altair	0.77	5.1	2.2	A7 V
α2 Aur	Capella B	0.96	13	0.4	G0 III
α Tau	Aldebaran	0.98	20	−0.5	K5 III
α Vir	Spica	1.04	77	−3.4	B1 III-IV
α Sco	Antares	1.09	170	−5.1	M1.5 Iab-b
β Gem	Pollux	1.15	10	1.2	K0 IIIb
α PsA	Fomalhaut	1.16	7.7	1.7	A4 V
α Cyg	Deneb	1.25	430	−6.9	A2 Iae
β Cru	Mimosa	1.30	85	−3.3	B0.5 IV
α Cen B	Rigil Kentaurus	1.33	1.3	5.8	K1 V
α Leo	Regulus	1.40	24	−0.5	B8 IV
α Cru A	Acrux	1.40	99	−3.6	B0.5 IV
ε CMa	Adara	1.51	120	−3.9	B2 Iab
λ Sco	Shaula	1.62	180	−4.7	B2 IV
γ Cru	Gacrux	1.63	27	−0.5	M3.5III
γ Ori	Bellatrix	1.64	77	−2.8	B2 III
β Tau	El Nath	1.68	41	−1.4	B7 III
β Car	Miaplacidus	1.70	35	−1.0	A2 IV
ε Ori	Alnilam	1.70	610	−7.2	B0 Iab
α Gru	Alnair	1.74	31	−0.7	B6 V
ε UMa	Alioth	1.76	25	−0.2	A0pCr
ζ Ori A	Alnitak	1.77	230	−5.0	O9 Iab
α UMa A	Dubhe	1.79	38	−1.1	K0 Iab
ε Sgr	Kaus Australis	1.80	44	−1.4	B9.5 III
γ2 Vel	Suhail	1.81	340	−5.8	O7.5e
α Per	Mirfak	1.82	160	−4.2	F5 Iab
δ CMa	Wezen	1.84	150	−4.0	F8 Iab
η UMa	Alkaid	1.85	32	−0.7	B3 V
θ Sco	Sargas	1.86	92	−3.0	F1 II
γ Gem	Alhena	1.90	34	−0.8	A0 IV
α Pav	Peacock	1.91	55	−1.8	B2 IV
α TrA	Atria	1.92	120	−3.5	K2 IIb-IIIa
α Gem A	Castor A	1.93	15	1.0	A1 V
δ Vel A	Koo She	1.95	25	0.0	A1 V

(*continued*)

Star	Name	Apparent Visual Magnitude (m_v)	Distance[3] (pc)	Absolute Visual Magnitude[4] (M_v)	Spectral Type
α Ari	Hamal	2.00	20	0.5	K2III
β CMa	Mirzam	2.00	150	−3.9	B1 II-III
α Hya	Alphard	2.00	55	−1.7	K3 II-III
α UMi	Polaris	2.00	130	−3.6	F7 Ib-II
β Cet	Deneb Kaitos	2.04	29	−0.3	K0 III

[1]Data from the SIMBAD database, operated at CDS, Strasbourg, France.
[2]For multiple star systems, the magnitude given is the combined light of all components; the spectral type is for the primary component.
[3]Rounded to two significant figures.
[4]Computed from m_v and distance.

Table A-9 | The Nearest Stars[1]

Name	Distance (pc)	Apparent Visual Magnitude (m_v)	Absolute Visual Magnitude (M_v)	Spectral Type
Sun		−26.7	4.85	G2V
Proxima Centauri	1.30	11.1	15.5	M6Ve
α Centauri A	1.33	0.01	4.40	G2V
α Centauri B	1.33	1.34	5.70	K1V
Barnard's Star	1.83	9.51	13.2	M4.0Ve
Wolf 359	2.39	13.44	16.6	M6.0V
Lalande 21185	2.54	7.47	10.4	M2.0V
Sirius A	2.63	−1.46	1.42	A1V
Sirius B	2.63	8.44	11.3	WD
Luyten 726-8	2.68	12.54	15.4	M5.5Ve
WISE 1541-2250	2.85	21.20	23.9	Y
Ross 154	2.97	10.43	13.1	M3Ve
Ross 248	3.16	12.29	14.8	M6.0Ve
Epsilon Eridan	3.22	3.73	6.20	K2V
Lacaille 9352	3.27	7.34	9.80	M2Ve
Ross 128	3.35	11.10	13.5	M4.5Vn
EZ Aquarii A	3.45	13.33	15.6	M5.0Ve
EZ Aquarii B	3.45	13.27	15.6	M?
EZ Aquarii C	3.45	14.03	16.3	M?
Procyon A	3.50	0.34	2.70	F5IV-V
Procyon B	3.50	10.70	13.0	WD
61 Cygni A	3.50	5.21	7.50	K5.0V
61 Cygni B	3.50	6.03	8.30	K7.0V
Struve 2398 A	3.57	8.90	11.2	M3.0V
Struve 2398 B	3.57	9.69	12.0	M3.5V
Groombridge 34 A	3.56	8.08	10.3	M2.0V
Groombridge 34 B	3.56	11.06	13.3	M6.0V
Epsilon Indi A	3.63	4.69	6.90	K5Ve
Epsilon Indi Ba	3.63	24.12	26.3	T1.0V
Epsilon Indi Bb	3.63	>23	>25	T6.0V
DX Cancri	3.63	14.78	17.0	M6.5Ve
Tau Ceti	3.64	3.49	5.70	G8.5Vp
GJ 1061	3.68	13.09	15.3	M5.5V

(*continued*)

Name	Distance (pc)	Apparent Visual Magnitude (m_v)	Absolute Visual Magnitude (M_v)	Spectral Type
YZ Ceti	3.72	12.02	14.2	M4V
Luyten's Star	3.79	9.86	12.0	M3.5V
Teegarden's star	3.84	15.40	17.2	M7V
SCR 1845-6357 A	3.85	17.39	19.4	M8.5V
SCR 1845-6357 B	3.85	?	?	T6
Kapteyn's Star	3.92	8.84	10.9	M1.0V
Lacaille 8760	3.95	6.67	8.70	M0V
UGPS 0722-05	4.07	16.52	18.5	T9
Kruger 60 A	4.03	9.79	11.8	M3.0V
Kruger 60 B	4.03	11.41	13.4	M4.0V
DEN 1048-3956	4.04	17.39	19.4	M9V
Ross 614 A	4.09	11.07	13.1	M4.5V
Ross 614 B	4.09	14.23	16.2	M8V
Wolf 1061	4.24	10.07	11.9	M3.5V
Van Maanen's star	4.31	12.38	14.2	WD
Gliese 1	4.36	8.55	10.4	M1.5V
Wolf 424 A	4.39	13.18	15.0	M4Ve

[1]Data from the SIMBAD database, operated at CDS, Strasbourg, Franc.

Table A-10 | Properties of the Planets

ORBITAL PROPERTIES

Planet	Semimajor Axis (*a*) (AU)	Semimajor Axis (*a*) (10^6 km)	Orbital Period (*P*) (y)	Orbital Period (*P*) (days)	Average Orbital Velocity (km/s)	Orbital Eccentricity	Inclination to Ecliptic
Mercury	0.387	57.9	0.241	88.0	47.9	0.206	7.0°
Venus	0.723	108	0.615	224.7	35.0	0.007	3.4°
Earth	1.00	150	1.00	365.3	29.8	0.017	0°*
Mars	1.52	228	1.88	687.0	24.1	0.093	1.8°
Jupiter	5.20	779	11.9	4332	13.1	0.049	1.3°
Saturn	9.58	1433	29.5	10,759	9.7	0.056	2.5°
Uranus	19.23	2877	84.3	30,799	6.8	0.044	0.8°
Neptune	30.10	4503	164.8	60,190	5.4	0.011	1.8°

* By definition.

PHYSICAL PROPERTIES (Earth = ⊕)

Planet	Equatorial Radius (km)	Equatorial Radius (⊕ = 1)	Mass (⊕ = 1)	Average Density (g/cm^3)	Surface Gravity (⊕ = 1)	Escape Velocity (km/s)	Sidereal Period of Rotation	Inclination of Equator to Orbit
Mercury	2440	0.383	0.055	5.43	0.38	4.2	58.6d	0.0°
Venus	6052	0.950	0.815	5.20	0.90	10.5	243.0d	177.3°
Earth	6378	1.00	1.000	5.52	1.00	11.2	23.93h	23.4°
Mars	3396	0.533	0.107	3.93	0.38	5.0	24.62h	25.2°
Jupiter	71,492	11.21	318	1.33	2.53	59.5	9.92h	3.1°
Saturn	60,268	9.45	95.2	0.69	1.06	35.5	10.57h	26.7°
Uranus	25,559	4.01	14.5	1.27	0.89	21.3	17.24h	97.8°
Neptune	24,764	3.88	17.1	1.64	1.14	23.5	16.11h	28.3°

Table A-11 | Principal Satellites of the Solar System

Planet	Satellite	Radius (km)	Distance from Planet (10^3 km)	Orbital Period (days)	Orbital Eccentricity	Orbital Inclination
Earth	Moon	1738	384.4	27.32	0.055	5.1°
Mars	Phobos	14×12×10	9.4	0.32	0.018	1.0°
	Deimos	8×6×5	23.5	1.26	0.002	2.8°
Jupiter	Amalthea	135×100×78	182	0.50	0.003	0.4°
	Io	1820	422	1.77	0.000	0.3°
	Europa	1565	671	3.55	0.000	0.5°
	Ganymede	2640	1071	7.16	0.002	0.2°
	Callisto	2420	1884	16.69	0.008	0.2°
	Himalia	~85*	11,470	250.6	0.158	27.6°
Saturn	Janus	110×80×100	151.5	0.70	0.007	0.1°
	Mimas	196	185.5	0.94	0.020	1.5°
	Enceladus	250	238.0	1.37	0.004	0.0°
	Tethys	530	294.7	1.89	0.000	1.1°
	Dione	560	377	2.74	0.002	0.0°
	Rhea	765	527	4.52	0.001	0.4°
	Titan	2575	1222	15.94	0.029	0.3°
	Hyperion	205×130×110	1484	21.28	0.104	~0.5°
	Iapetus	720	3562	79.33	0.028	14.7°
	Phoebe	110	12,930	550.4	0.163	150°
Uranus	Miranda	242	129.9	1.41	0.017	3.4°
	Ariel	580	190.9	2.52	0.003	0°
	Umbriel	595	266.0	4.14	0.003	0°
	Titania	805	436.3	8.71	0.002	0°
	Oberon	775	583.4	13.46	0.001	0°
Neptune	Proteus	205	117.6	1.12	~0	~0°
	Triton	1352	354.59	5.88	0.00	160°
	Nereid	170	5588.6	360.12	0.76	27.7°

*The ~ symbol means "approximately."

Table A-12 | Meteor Showers

Shower	Dates	Hourly Rate	Radiant of shower R.A.	Radiant of shower Dec.	Associated Comet
Quadrantids	Jan. 2–4	30	15h24m	50°	
Lyrids	April 20–22	8	18h4m	33°	Thatcher
η Aquarids	May 2–7	10	22h24m	0°	Halley
δ Aquarids	July 26–31	15	22h36m	−10°	
Perseids	Aug. 10–14	40	3h4m	58°	Swift-Tuttle
Orionids	Oct. 18–23	15	6h20m	15°	Halley
Taurids	Nov. 1–7	8	3h40m	17°	Encke
Leonids	Nov. 14–19	6	10h12m	22°	Tempel-Tuttle
Geminids	Dec. 10–13	50	7h28m	32°	

Table A-13 | Greatest Elongations of Mercury*

Evening Sky	Morning Sky
Oct. 26, 2012	Dec. 4, 2012
Feb. 16, 2013	March 31, 2013
June 12, 2013	July 30, 2013
Oct. 9, 2013	Nov. 18, 2013
Jan. 31, 2014	March 14, 2014
May 25, 2014	July 12, 2014
Sept. 21, 2014	Nov. 1, 2014**
Jan. 14, 2015	Feb. 24, 2015
May 7, 2015	June 24, 2015
Sept. 4, 2015	Oct. 16, 2015**
Dec. 29, 2015`	Feb. 7, 2016

*Elongation is the angular distance from the sun to a planet.
**Most favorable elongations.

Table A-14 | Greatest Elongations of Venus[1]

Evening Sky	Morning Sky
Nov. 1, 2013	March 22, 2014
June 6, 2015	Oct. 26, 2015
Jan. 12, 2017	June 3, 2017
Aug. 17, 2018	Jan. 6, 2019
March 24, 2020	Aug. 12, 2020
Oct. 29, 2021	March 20, 2022
June 4, 2023	Oct. 23, 2023

[1]Venus does not come to elongation during 2024.

Table A-15 | The Greek Alphabet

A, α	alpha	H, η	eta	N, ν	nu	T, τ	tau
B, β	beta	Θ, θ	theta	Ξ, ξ	xi	Υ, υ	upsilon
Γ, γ	gamma	I, ι	iota	O, o	omicron	Φ, ϕ	phi
Δ, δ	delta	K, κ	kappa	Π, π	pi	X, χ	chi
E, ε	epsilon	Λ, λ	lambda	P, ρ	rho	Ψ, ψ	psi
Z, ζ	zeta	M, μ	mu	Σ, σ	sigma	Ω, ω	omega

■ **Table A-16** | **Periodic Table of the Elements**

Atomic number → 11
Symbol → Na
Atomic mass → 22.99

Atomic masses are based on carbon-12. Numbers in parentheses are mass numbers of most stable or best-known isotopes of radioactive elements.

Group

Period	IA(1)	IIA(2)	IIIB(3)	IVB(4)	VB(5)	VIB(6)	VIIB(7)	VIII (8)	VIII (9)	VIII (10)	IB(11)	IIB(12)	IIIA(13)	IVA(14)	VA(15)	VIA(16)	VIIA(17)	Noble Gases (18)
1	1 H 1.008																	2 He 4.003
2	3 Li 6.941	4 Be 9.012											5 B 10.81	6 C 12.01	7 N 14.01	8 O 16.00	9 F 19.00	10 Ne 20.18
3	11 Na 22.99	12 Mg 24.31											13 Al 26.98	14 Si 28.09	15 P 30.97	16 S 32.06	17 Cl 35.45	18 Ar 39.95
4	19 K 39.10	20 Ca 40.08	21 Sc 44.96	22 Ti 47.90	23 V 50.94	24 Cr 52.00	25 Mn 54.94	26 Fe 55.85	27 Co 58.93	28 Ni 58.7	29 Cu 63.55	30 Zn 65.38	31 Ga 69.72	32 Ge 72.59	33 As 74.92	34 Se 78.96	35 Br 79.90	36 Kr 83.80
5	37 Rb 85.47	38 Sr 87.62	39 Y 88.91	40 Zr 91.22	41 Nb 92.91	42 Mo 95.94	43 Tc 98.91	44 Ru 101.1	45 Rh 102.9	46 Pd 106.4	47 Ag 107.9	48 Cd 112.4	49 In 114.8	50 Sn 118.7	51 Sb 121.8	52 Te 127.6	53 I 126.9	54 Xe 131.3
6	55 Cs 132.9	56 Ba 137.3	57* La 138.9	72 Hf 178.5	73 Ta 180.9	74 W 183.9	75 Re 186.2	76 Os 190.2	77 Ir 192.2	78 Pt 195.1	79 Au 197.0	80 Hg 200.6	81 Tl 204.4	82 Pb 207.2	83 Bi 209.0	84 Po (210)	85 At (210)	86 Rn (222)
7	87 Fr (223)	88 Ra 226.0	89** Ac (227)	104 Rf (261)	105 Db (262)	106 Sg (263)	107 Bh (262)	108 Hs (265)	109 Mt (266)	110 Ds (269)	111 Uuu (272)	112 Uub (277)	113 Uub (284)	114 Uuq (285)	115 Uub (288)	116 Uuh (289)		

Transition Elements (groups 3–12)

Inner Transition Elements

Lanthanide Series 6 *	58 Ce 140.1	59 Pr 140.9	60 Nd 144.2	61 Pm (145)	62 Sm 150.4	63 Eu 152.0	64 Gd 157.3	65 Tb 158.9	66 Dy 162.5	67 Ho 164.9	68 Er 167.3	69 Tm 168.9	70 Yb 173.0	71 Lu 175.0
Actinide Series 7 **	90 Th 232.0	91 Pa 231.0	92 U 238.0	93 Np 237.0	94 Pu (244)	95 Am (243)	96 Cm (247)	97 Bk (247)	98 Cf (251)	99 Es (252)	100 Fm (257)	101 Md (258)	102 No (259)	103 Lr (260)

The Elements and Their Symbols

Actinium	Ac
Aluminum	Al
Americium	Am
Antimony	Sb
Argon	Ar
Arsenic	As
Astatine	At
Barium	Ba
Berkelium	Bk
Beryllium	Be
Bismuth	Bi
Bohrium	Bh
Boron	B
Bromine	Br
Cadmium	Cd
Calcium	Ca
Californium	Cf
Carbon	C
Cerium	Ce
Cesium	Cs
Chlorine	Cl
Chromium	Cr
Cobalt	Co
Copper	Cu
Curium	Cm
Darmstadtium	Ds
Dubnium	Db
Dysprosium	Dy
Einsteinium	Es
Erbium	Er
Europium	Eu
Fermium	Fm
Fluorine	F
Francium	Fr
Gadolinium	Gd
Gallium	Ga
Germanium	Ge
Gold	Au
Hafnium	Hf
Hassium	Hs
Helium	He
Holmium	Ho
Hydrogen	H
Indium	In
Iodine	I
Iridium	Ir
Iron	Fe
Krypton	Kr
Lanthanum	La
Lawrencium	Lr
Lead	Pb
Lithium	Li
Lutetium	Lu
Magnesium	Mg
Manganese	Mn
Meitnerium	Mt
Mendelevium	Md
Mercury	Hg
Molybdenum	Mo
Neodymium	Nd
Neon	Ne
Neptunium	Np
Nickel	Ni
Niobium	Nb
Nitrogen	N
Nobelium	No
Osmium	Os
Oxygen	O
Palladium	Pd
Phosphorous	P
Platinum	Pt
Plutonium	Pu
Polonium	Po
Potassium	K
Praseodymium	Pr
Promethium	Pm
Protactinium	Pa
Radium	Ra
Radon	Rn
Rhenium	Re
Rhodium	Rh
Rubidium	Rb
Ruthenium	Ru
Rutherfordium	Rf
Samarium	Sm
Scandium	Sc
Seaborgium	Sg
Selenium	Se
Silicon	Si
Silver	Ag
Sodium	Na
Strontium	Sr
Sulfur	S
Tantalum	Ta
Technetium	Tc
Tellurium	Te
Terbium	Tb
Thallium	Tl
Thorium	Th
Thulium	Tm
Tin	Sn
Titanium	Ti
Tungsten	W
Uranium	U
Vanadium	V
Xenon	Xe
Ytterbium	Yb
Yttrium	Y
Zinc	Zn
Zirconium	Zr

Appendix B
Observing the Sky

Observing the sky with the naked eye is as important to modern astronomy as picking up pretty pebbles is to modern geology. The sky is an inspiring natural wonder unimaginably bigger than the Grand Canyon, the Rocky Mountains, or any other site that tourists visit every year. To neglect the beauty of the sky is equivalent to geologists neglecting the beauty of the minerals they study. This supplement is meant to act as a tourist's guide to the sky. You analyzed the universe in the textbook's chapters; here you can admire it.

The brighter stars in the sky are visible even from the centers of cities with their air and light pollution. But in the countryside, only a few miles beyond the cities, the night sky is a velvety blackness strewn with thousands of glittering stars. From a wilderness location, far from the city's glare, and especially from high mountains, the night sky is spectacular.

Because of the rotation and orbital motion of Earth, you need more than one star chart to map the sky. Which chart you select depends on the month and the time of night.

Two sets of charts are included for two typical locations on Earth. The Northern Hemisphere star charts show the sky as seen from a northern latitude typical for the United States and Central Europe. The Southern Hemisphere star charts are appropriate for readers in Earth's Southern Hemisphere, including Australia, southern South America, and southern Africa.

To use the charts, select the appropriate chart and hold it overhead as shown in ■ Figure B-1. If you face south, turn the chart until the words *southern horizon* are at the bottom of the chart. If you face other directions, turn the chart appropriately.

Using Star Charts

The constellations are a fascinating cultural heritage of our planet, but they are sometimes a bit difficult to learn because of Earth's motion. The constellations above the horizon change with the time of night and the seasons.

Because Earth rotates eastward, the sky appears to rotate westward around Earth. A constellation visible overhead soon after sunset will appear to move westward, and in a few hours it will disappear below the horizon. Other constellations will rise in the east, so the sky changes gradually through the night.

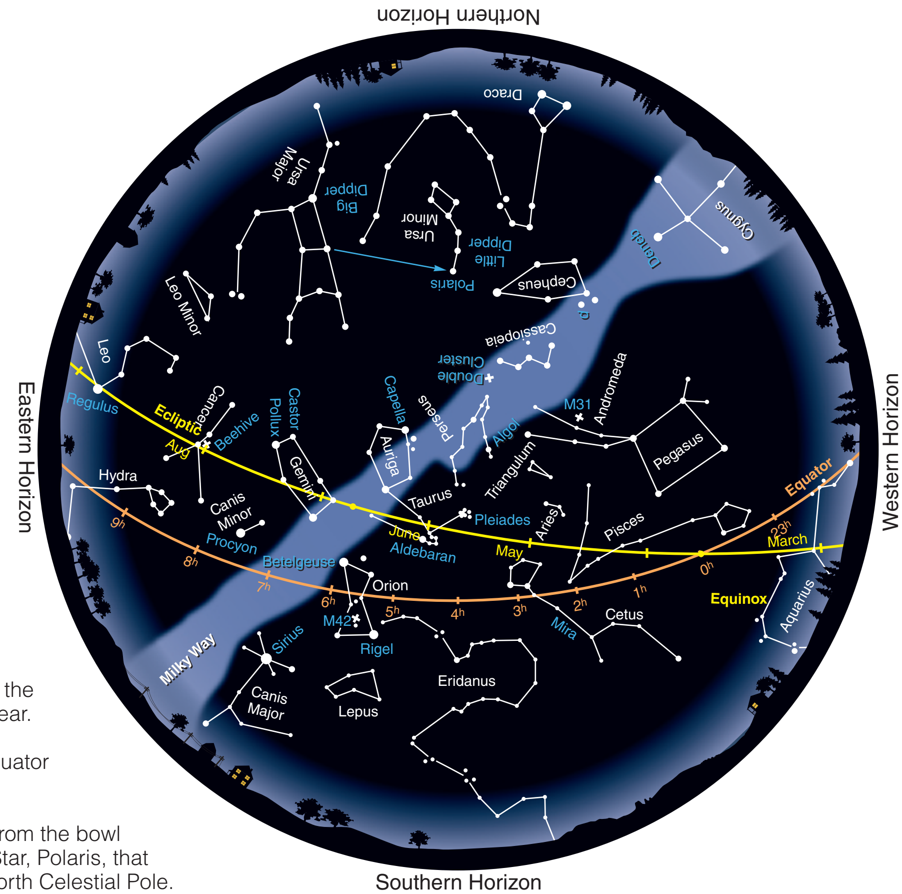
In addition, Earth's orbital motion makes the sun appear to move eastward among the stars. Each day the sun moves about twice its own diameter, or about one degree, eastward along the ecliptic; consequently, each night at sunset, the constellations are shifted about one degree farther toward the west.

Orion, for instance, is visible in the evening sky in January, but as the days pass, the sun moves closer to Orion. By March, Orion is difficult to see in the western sky soon after sunset. By June, the sun is so close to Orion that it sets with the sun and is invisible. Not until late July is the sun far enough past Orion for the constellation to become visible rising in the eastern sky just before dawn.

© Cengage Learning 2014

■ **Figure B-1**

To use the star charts in this book, select the appropriate chart for the season and time. Hold it overhead and turn it until the direction at the bottom of the chart is the same as the direction you are facing.



Northern Hemisphere Sky

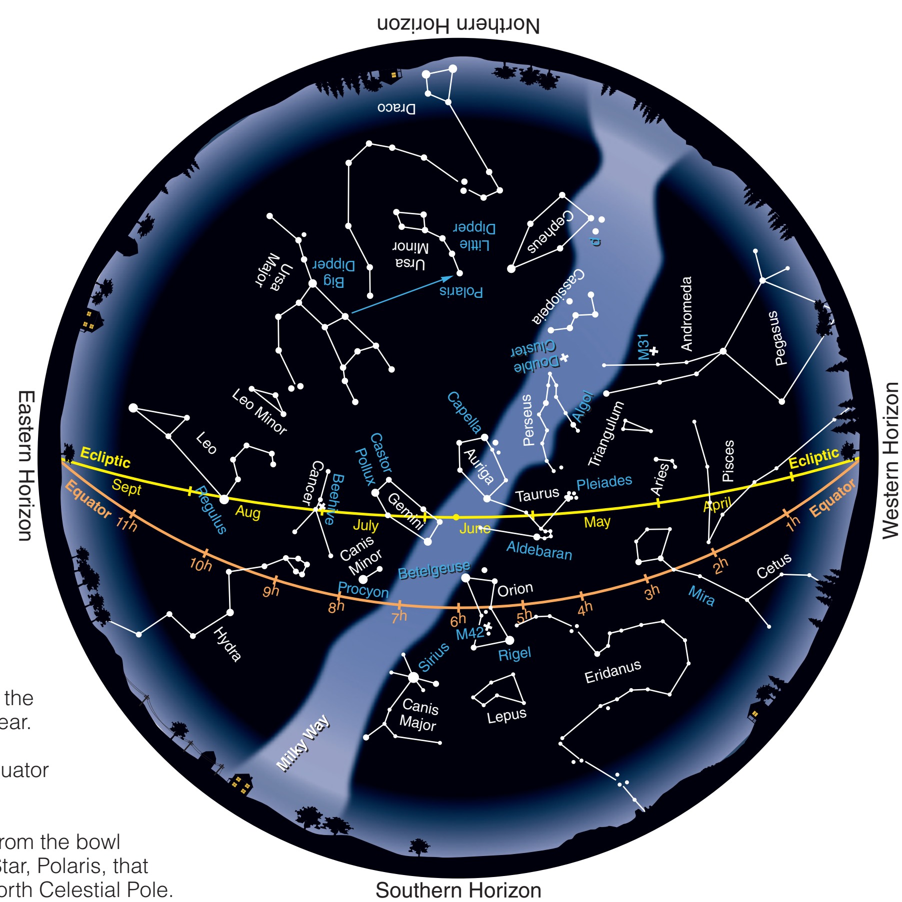
JANUARY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Northern Hemisphere Sky

FEBRUARY

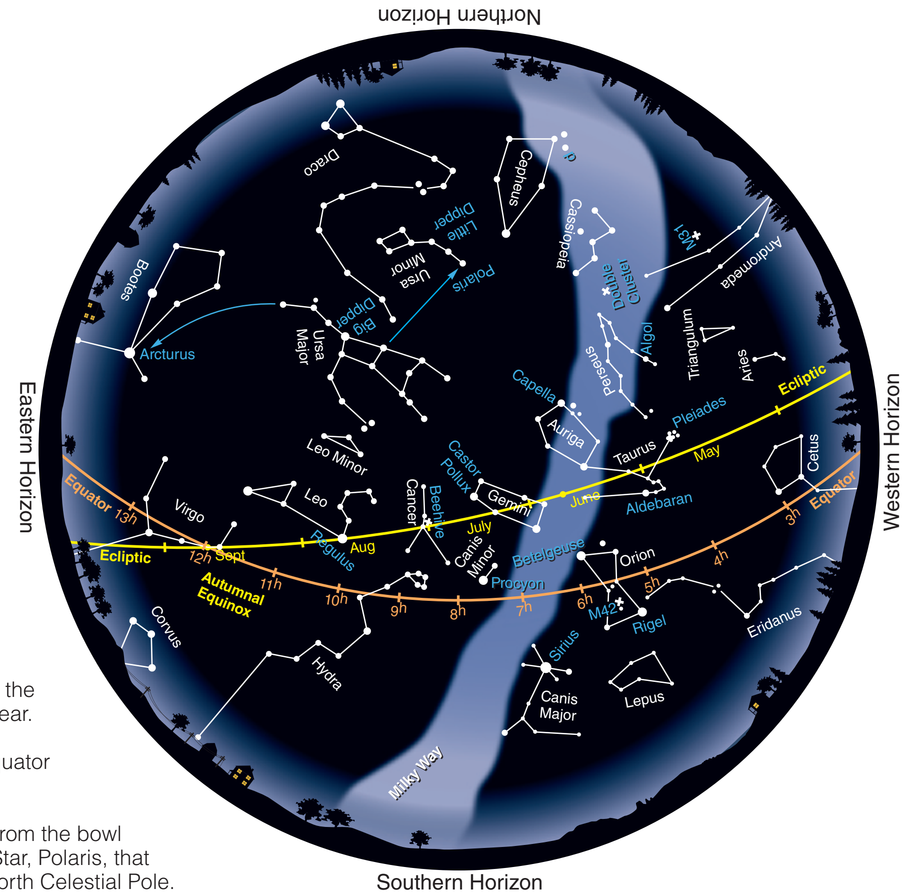
Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.

Line art on this page © Cengage Learning 2014



Northern Hemisphere Sky

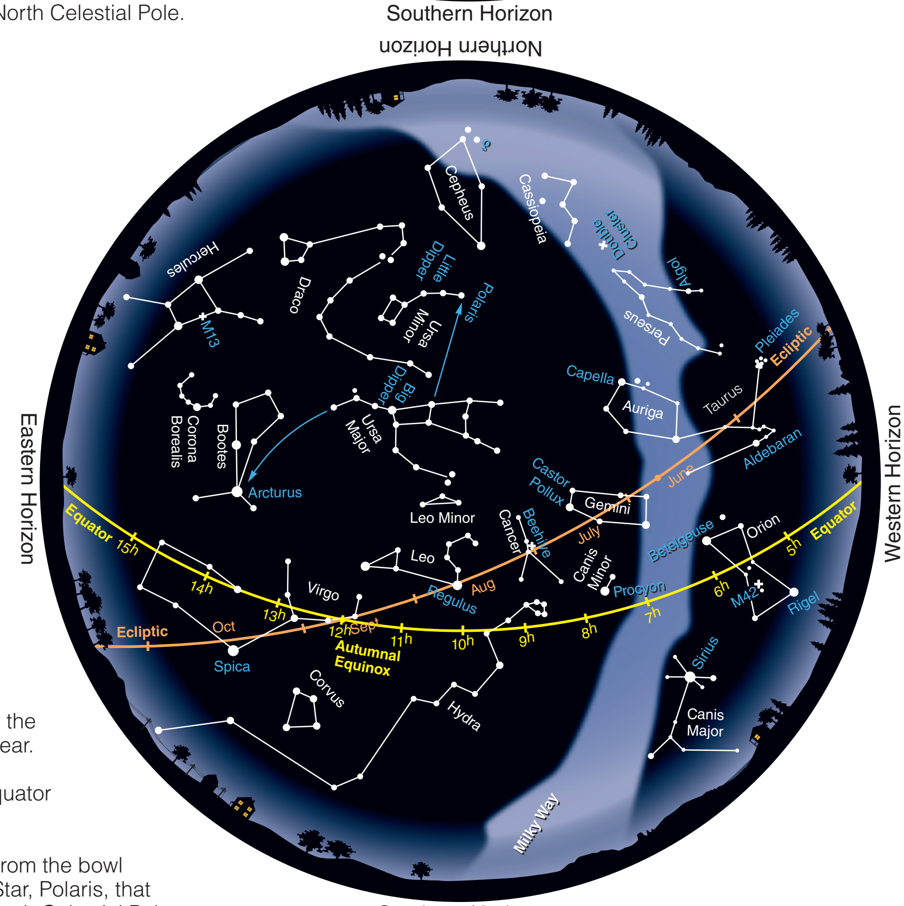
MARCH

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Northern Hemisphere Sky

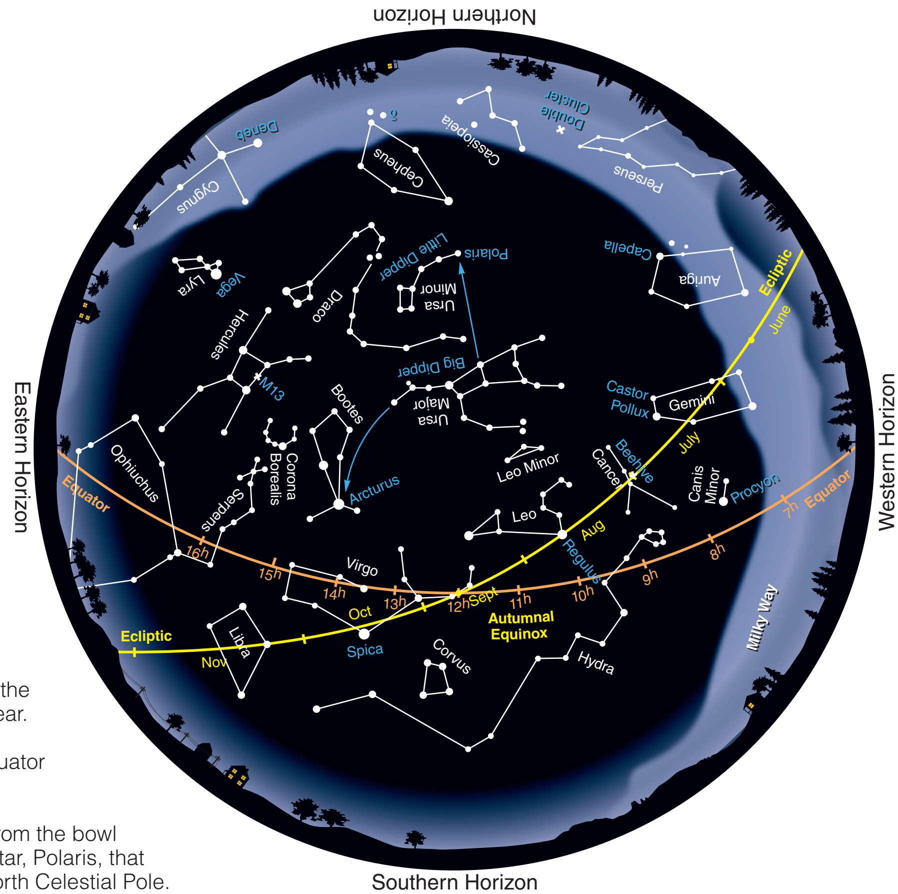
APRIL

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Northern Hemisphere Sky

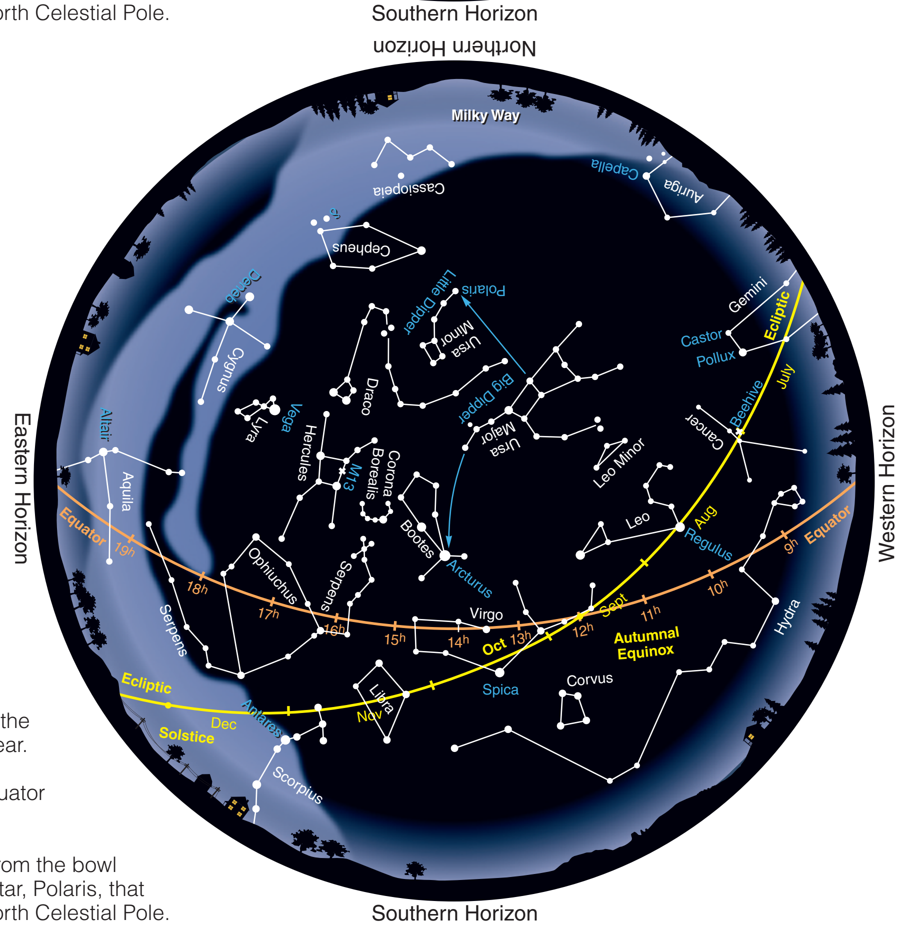
MAY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Northern Hemisphere Sky

JUNE

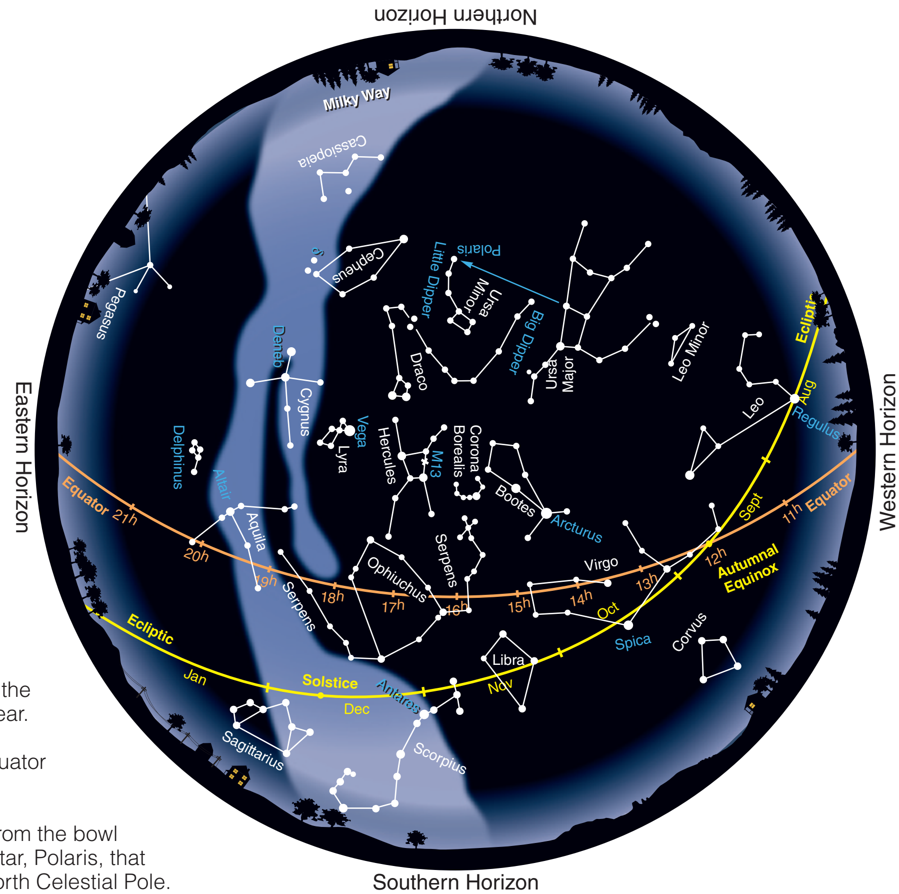
Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.

Line art on this page © Cengage Learning 2014



Northern Hemisphere Sky

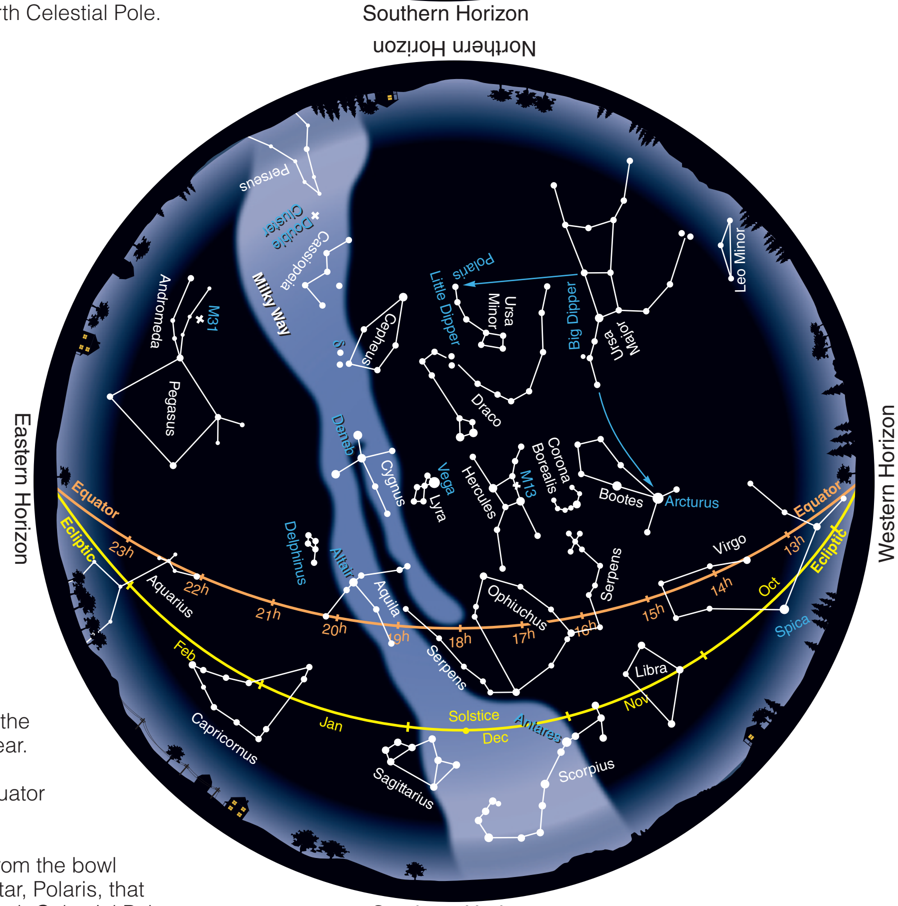
JULY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Northern Hemisphere Sky

AUGUST

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.

Line art on this page © Cengage Learning 2014

Northern Hemisphere Sky

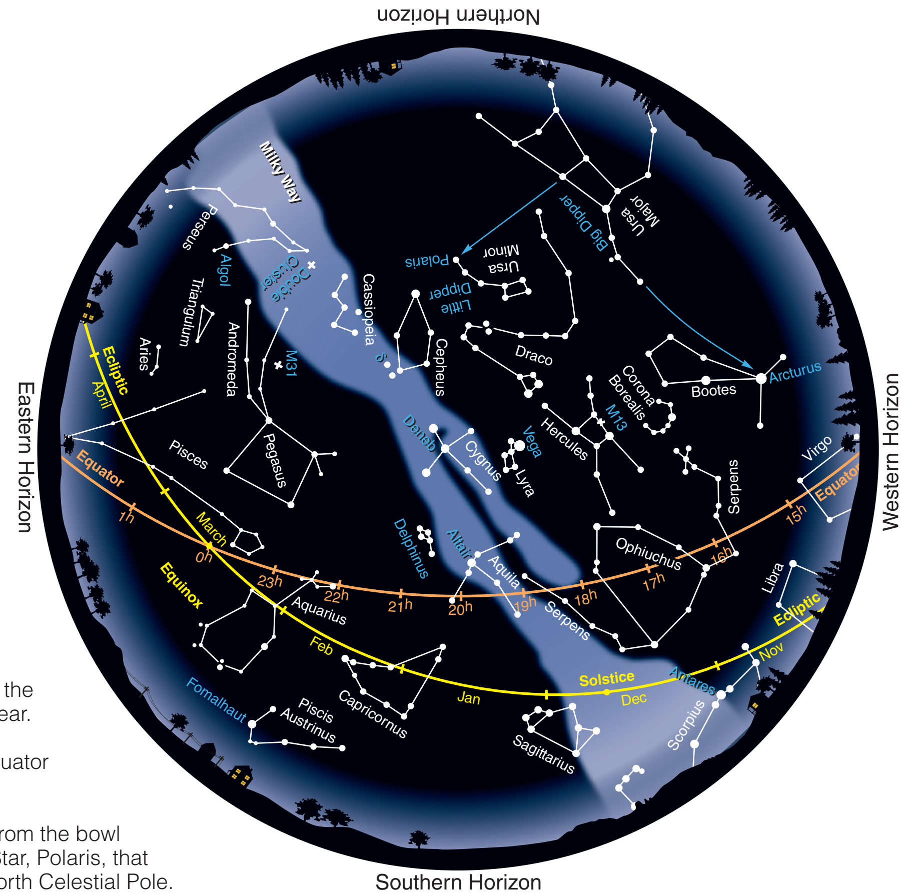
SEPTEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Northern Hemisphere Sky

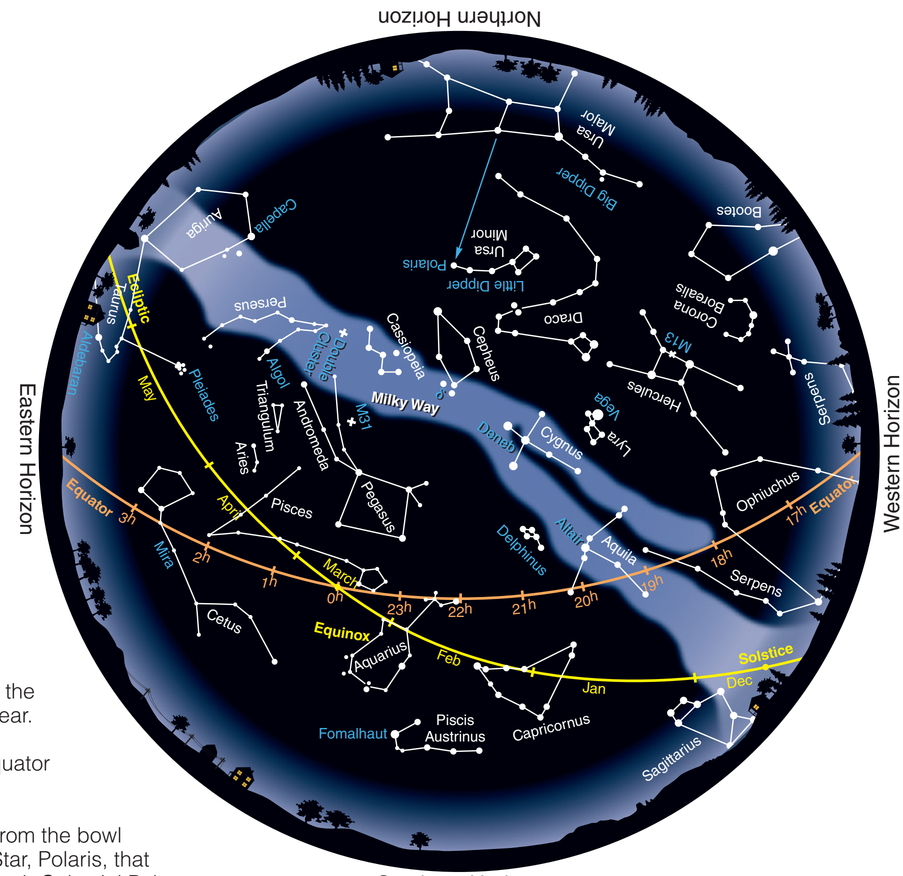
OCTOBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Line art on this page © Cengage Learning 2014

Northern Hemisphere Sky

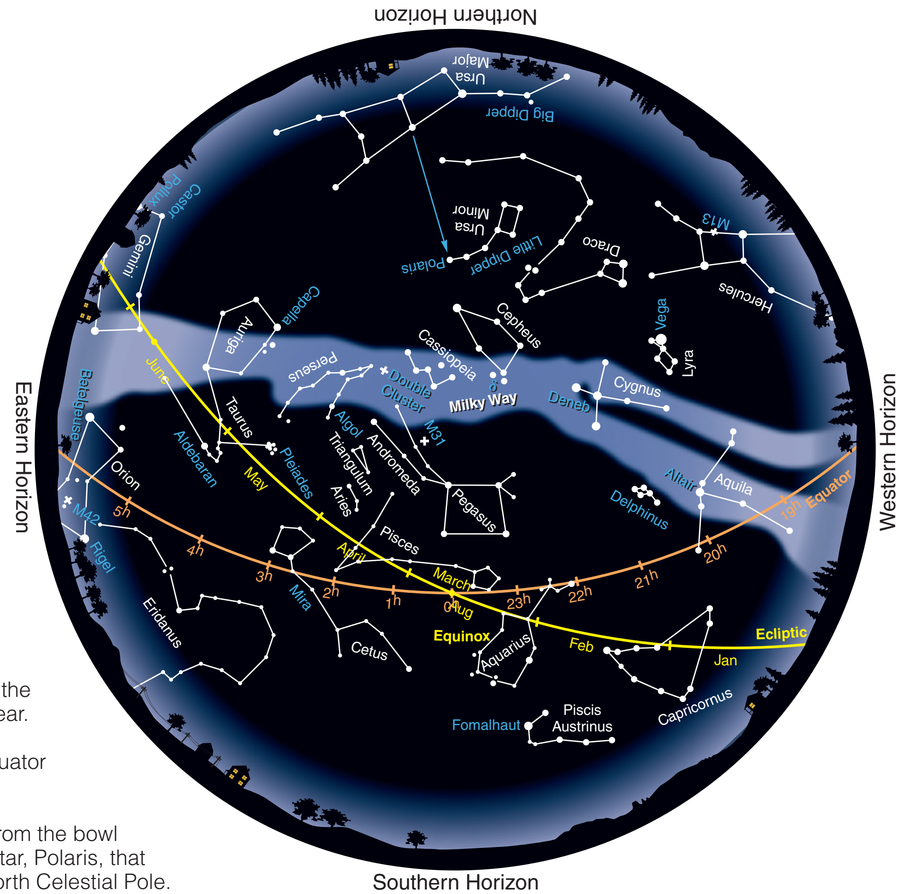
NOVEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Northern Hemisphere Sky

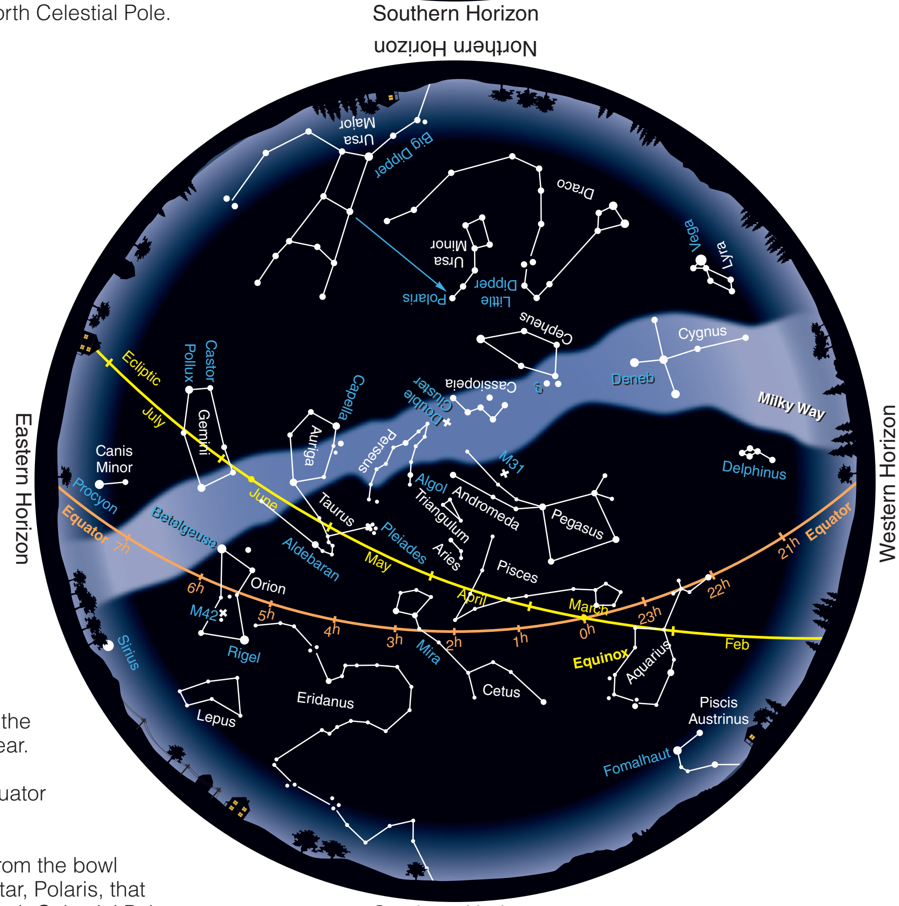
DECEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

The straight blue arrow points from the bowl of the Big Dipper to the North Star, Polaris, that is close to the location of the North Celestial Pole.



Line art on this page © Cengage Learning 2014

Southern Hemisphere Sky

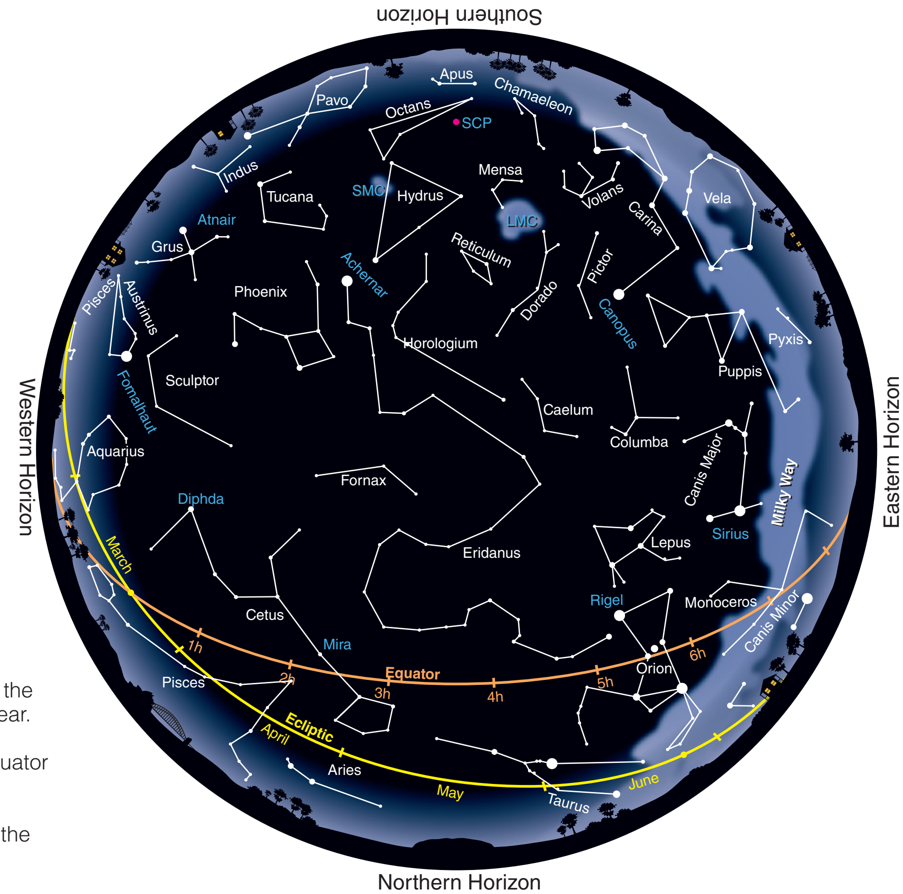
JANUARY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).



Southern Hemisphere Sky

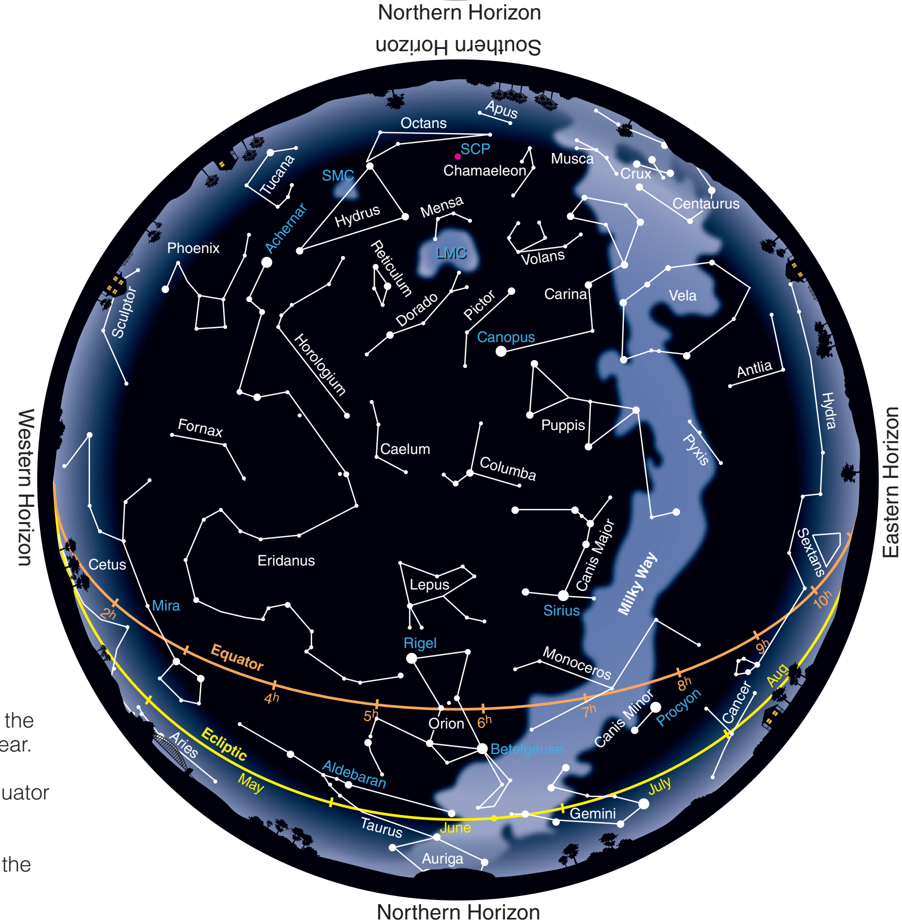
FEBRUARY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).



Line art on this page © Cengage Learning 2014

Southern Hemisphere Sky

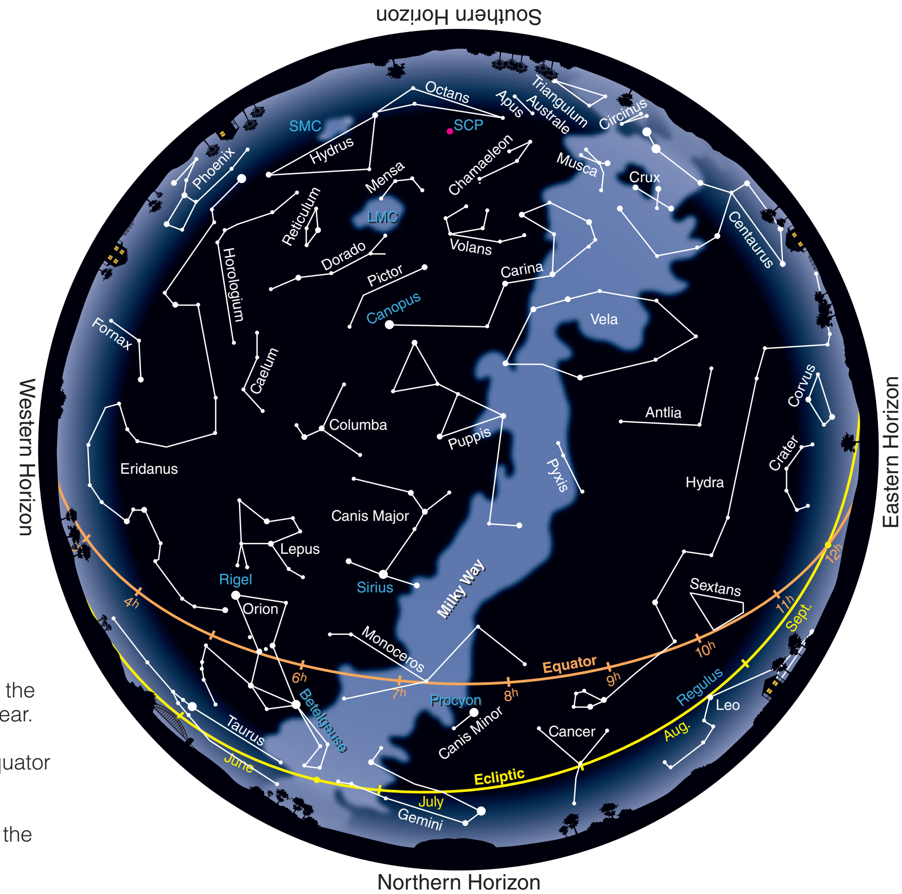
MARCH

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).



Southern Hemisphere Sky

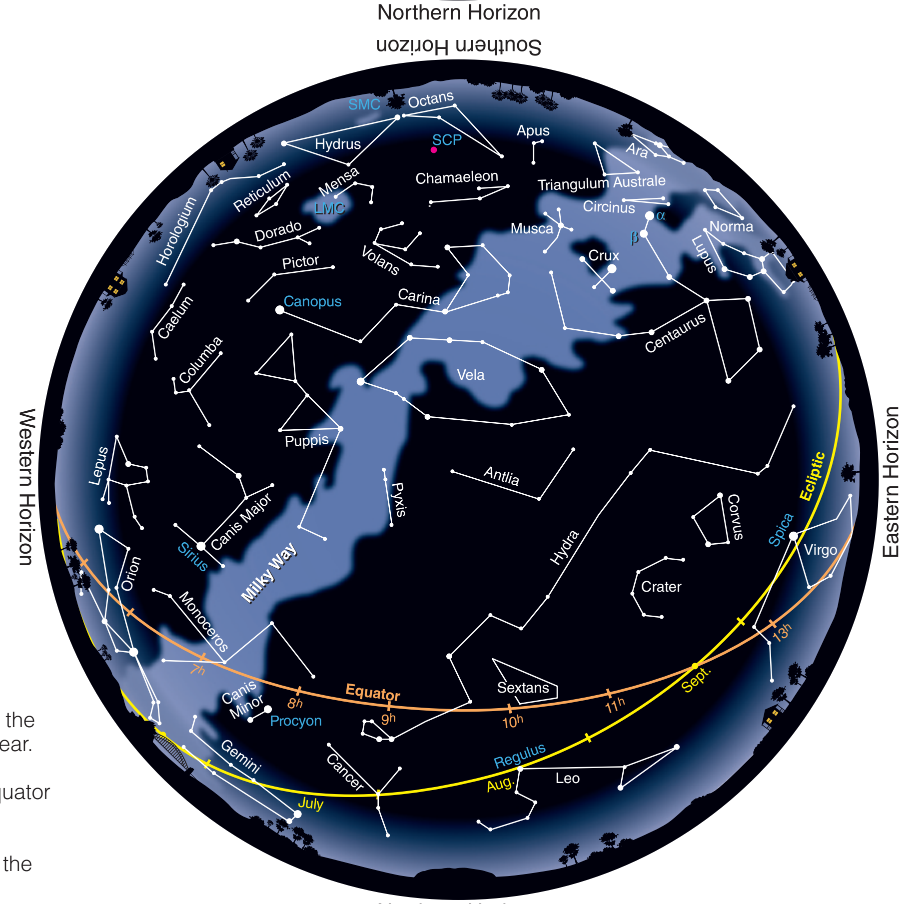
APRIL

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

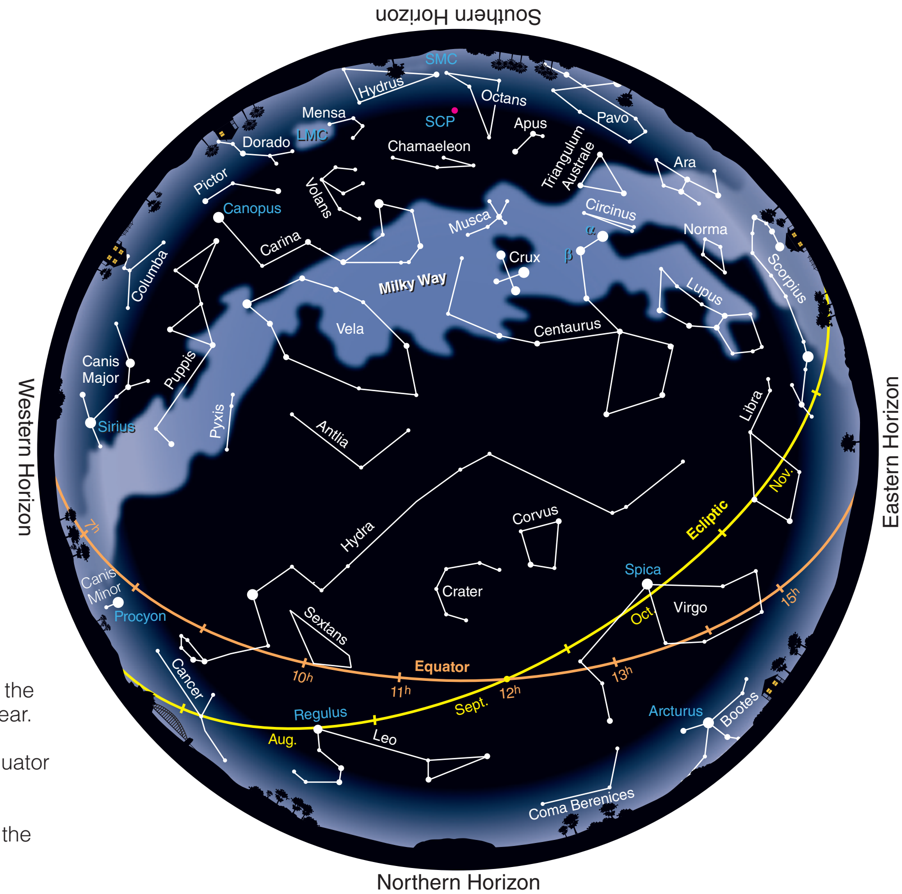
Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).



Line art on this page © Cengage Learning 2014



Southern Hemisphere Sky

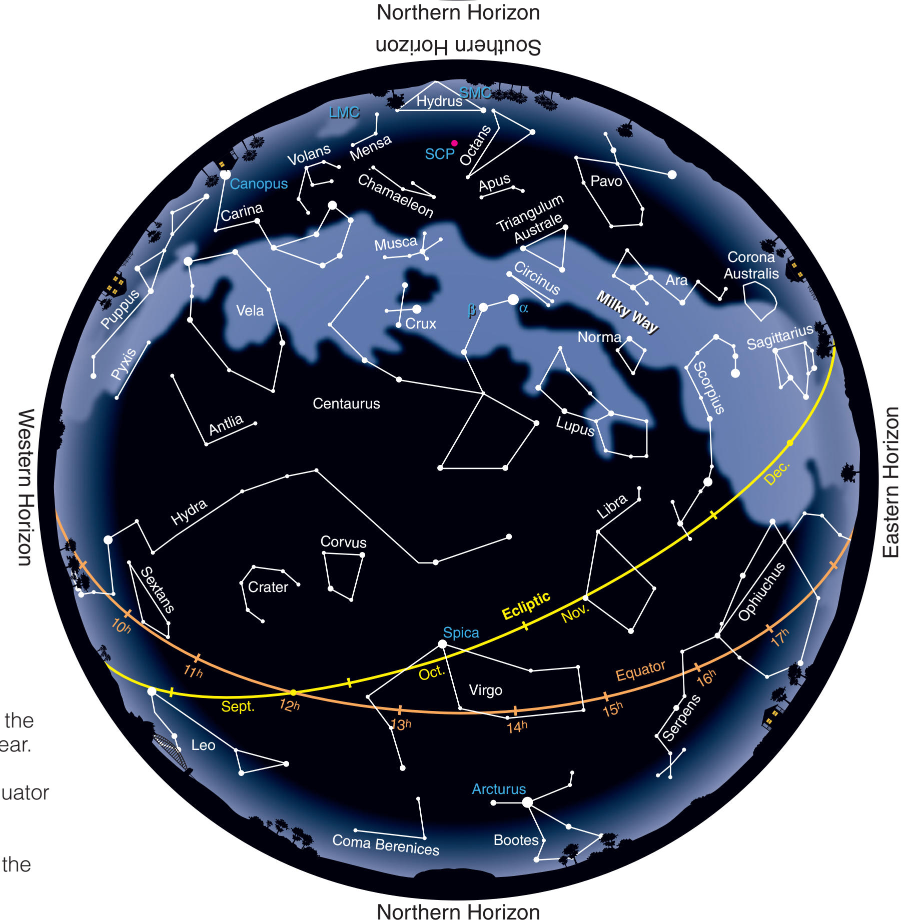
MAY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).



Southern Hemisphere Sky

JUNE

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).

Southern Hemisphere Sky

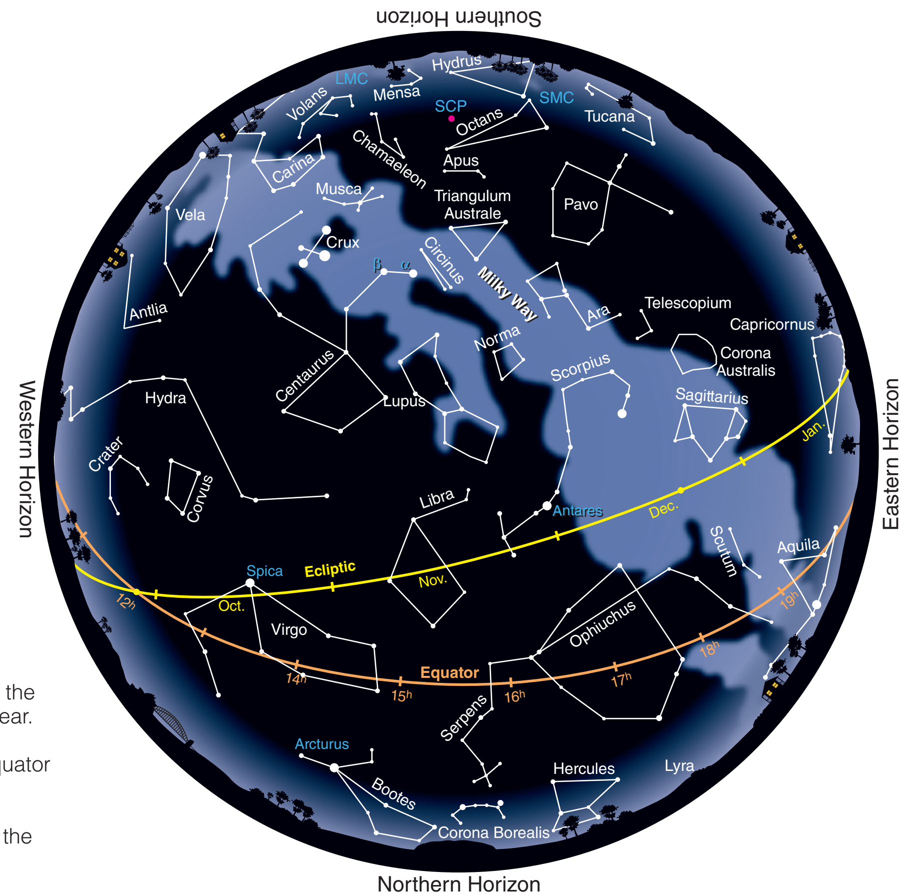
JULY

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).



Southern Hemisphere Sky

AUGUST

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).

Southern Hemisphere Sky

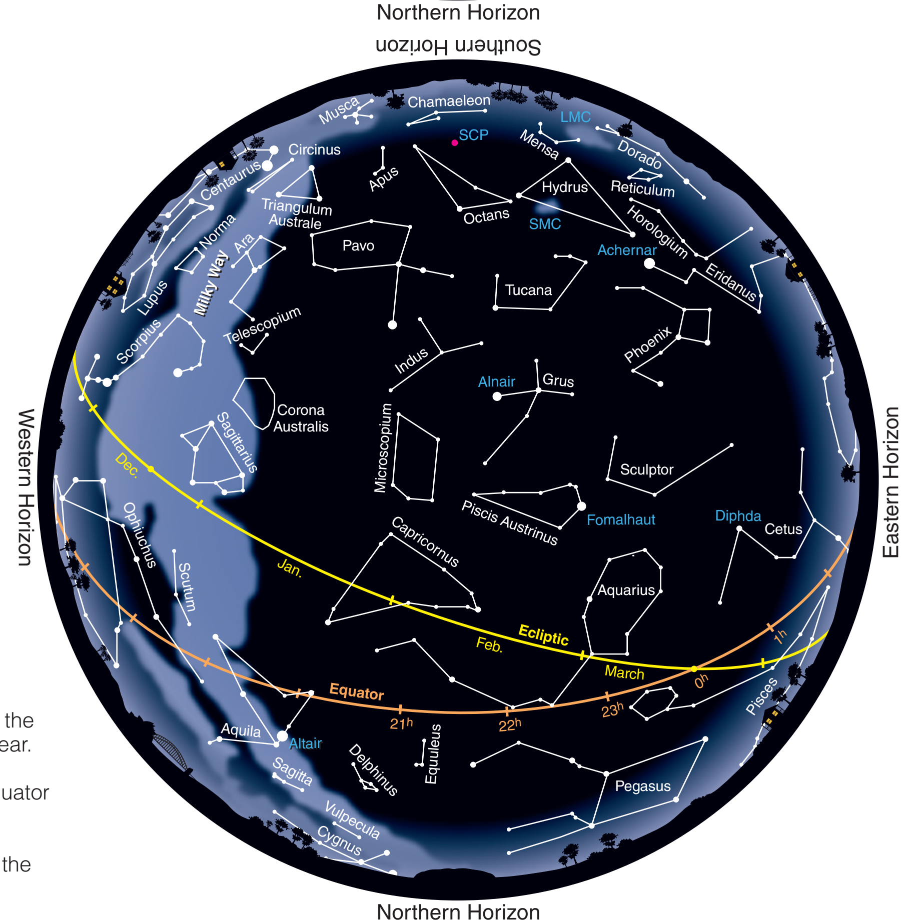
SEPTEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).



Southern Hemisphere Sky

OCTOBER

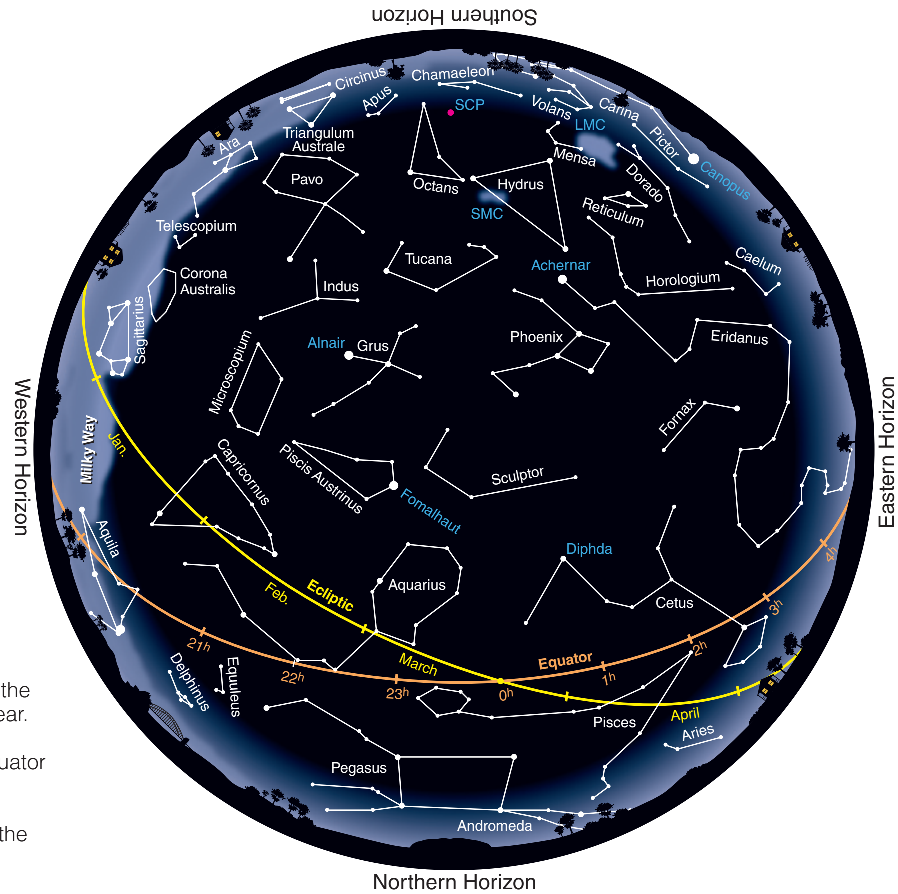
Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).

Line art on this page © Cengage Learning 2014



Southern Hemisphere Sky

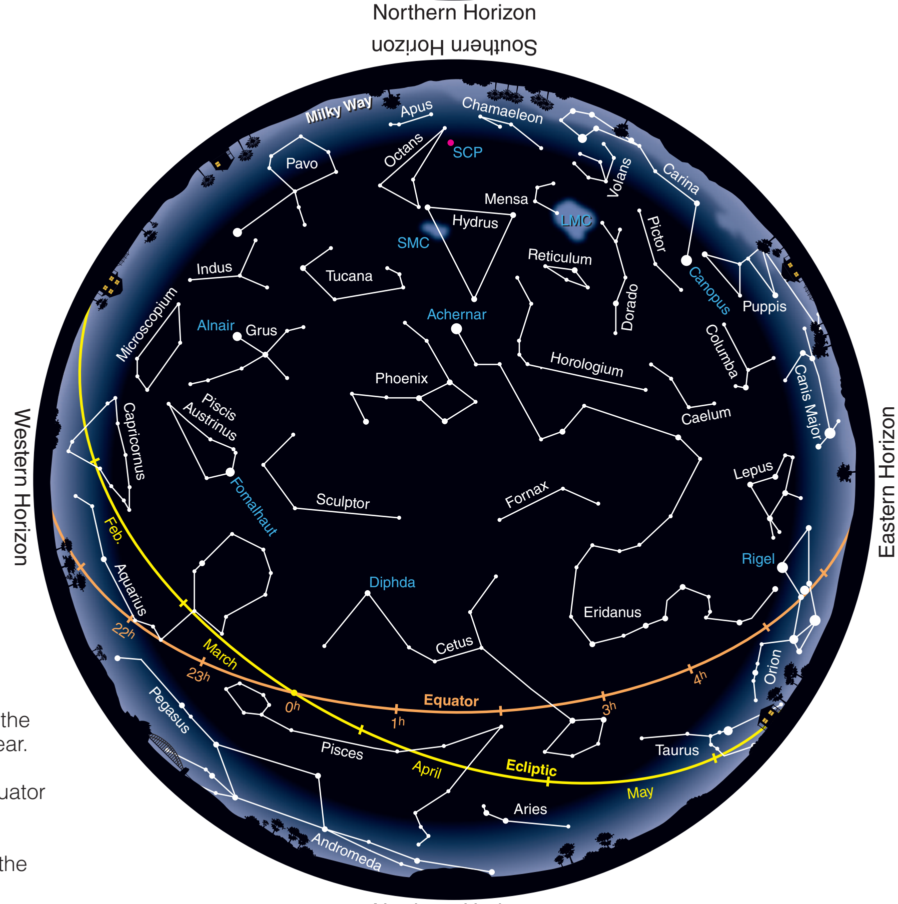
NOVEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).



Southern Hemisphere Sky

DECEMBER

Early in Month	9 P.M.
Midmonth	8 P.M.
End of Month	7 P.M.

Months along the ecliptic show the location of the sun during the year.

Numbers along the celestial equator show right ascension.

A red dot marks the location of the South Celestial Pole (SCP).

Line art on this page © Cengage Learning 2014

Glossary

Pronunciation Guide

ā pay	ē be	ī pie	ō so	ŭ cut
ă hat	ĕ pet	ĭ pit	ŏ pot	
ä father		î pier	ô paw	
			o͞o food	

Numbers in parentheses refer to the page where the term is first discussed in the text.

absolute visual magnitude (M_V) (262) Intrinsic brightness of a star; the apparent visual magnitude the star would have if it were 10 pc away.

absolute zero (107) The lowest possible temperature; the temperature at which the particles in a material, atoms or molecules, contain no energy of motion that can be extracted from the body.

absorption line (110) A dark line in a spectrum; produced by the absence of photons absorbed by atoms or molecules.

absorption (dark-line) spectrum (110) A spectrum that contains absorption lines.

accretion (152) The sticking together of solid particles to produce a larger particle.

accretion disk (332) The whirling disk of gas that forms around a compact object such as a white dwarf, neutron star, or black hole as matter is drawn in.

achondrite (241) Stony meteorite containing no chondrules or volatiles.

achromatic lens (78) A telescope lens composed of two lenses ground from different kinds of glass and designed to bring two selected colors to the same focus and correct for chromatic aberration.

active galactic nucleus (AGN) (409) The central energy source of an active galaxy.

active galaxy (409) A galaxy that is a source of excess radiation, usually radio waves, X-rays, gamma rays, or some combination.

active optics (89) Optical elements whose position or shape is continuously controlled by computers.

active region (127) An area on the sun where sunspots, prominences, flares, and the like occur.

adaptive optics (95) Computer-controlled telescope mirrors that can at least partially compensate for seeing.

albedo (174) The fraction of the light hitting an object that is reflected.

alt-azimuth mount (89) A telescope mounting capable of motion parallel to and perpendicular to the horizon.

amino acid (456) One of the carbon-chain molecules that are the building blocks of protein.

angstrom (Å) (76) A unit of distance; 1 Å $= 10^{-10}$ m; often used to measure the wavelength of light.

angular diameter (19) A measure of the size of an object in the sky; numerically equal to the angle in degrees between two lines extending from the observer's eye to opposite edges of the object.

angular distance (19) A measure of the separation between two objects in the sky; numerically equal to the angle in degrees between two lines extending from the observer's eye to the two objects.

angular momentum (332) The tendency of a rotating body to continue rotating; mathematically, the product of mass, velocity, and radius.

annular eclipse (39) A solar eclipse in which the solar photosphere appears around the edge of the moon in a bright ring, or annulus. The corona, chromosphere, and prominences cannot be seen.

anorthosite (177) Rock of aluminum and calcium silicates found in the lunar highlands.

anticoded (466) Message designed to be understood by a recipient about whom the sender knows little or nothing, for example an interstellar broadcast aimed at possible inhabitants of another planet.

antimatter (431) Matter composed of antiparticles, which on colliding with a matching particle of normal matter annihilate and convert the mass of both particles into energy. The antiproton is the antiparticle of the proton, and the positron is the antiparticle of the electron.

aphelion (27) The orbital point of greatest distance from the sun.

apogee (39) The orbital point of greatest distance from Earth.

Apollo object (246) Asteroid whose orbit crosses that of Earth (Apollo) and Mars (Amor).

apparent visual magnitude, m_v (15) The brightness of a star as seen by human eyes on Earth.

arc minute (19) An angular measure; each degree is divided into 60 arc minutes.

arc second (19) An angular measure; each arc minute is divided into 60 arc seconds.

array detector (93) A grid of photosensitive detectors for the purpose of recording images. Commercial still and video digital cameras contain CCD array detectors. See also *charge-coupled device.*

association (303) Group of widely scattered stars (10 to 1000) moving together through space; not gravitationally bound into a cluster.

asterism (12) A named group of stars not identified as a constellation, for example, the Big Dipper.

asteroid (145) Small, rocky world; most asteroids lie between Mars and Jupiter in the asteroid belt.

astrobiology (454) The field of study involving searches for life on other worlds and investigation of possible habitats for such life. Also known as "exobiology."

astronomical unit (AU) (3) Average distance from Earth to the sun; 1.5×10^8 km, or 93×10^6 miles.

astrophysics (101) The application of physics to the study of astronomical objects and the entire universe.

atmospheric window (77) Wavelength region in which Earth's atmosphere is transparent—at visual, infrared, and radio wavelengths.

aurora (133) The glowing light display that results when a planet's magnetic field guides charged particles toward the north and south magnetic poles, where they strike the upper atmosphere and excite atoms to emit photons.

autumnal equinox (26) The point on the celestial sphere where the sun crosses the celestial equator going southward. Also, the time when the sun reaches this point and autumn begins in the Northern Hemisphere—about September 22.

Babcock model (128) A model of the sun's magnetic cycle in which the differential rotation of the sun winds up and tangles the solar magnetic field in a 22-year cycle. This is thought to be responsible for the 11-year sunspot cycle.

Balmer series (111) Spectral lines in the visible and near-ultraviolet spectrum of hydrogen produced by transitions whose lowest orbit is the second.

barred spiral galaxy (398) A spiral galaxy with an elongated nucleus resembling a bar from which the arms originate.

basalt (172) Dark, igneous rock characteristic of solidified lava.

belt–zone circulation (211) The atmospheric circulation typical of Jovian planets. Dark belts and bright zones encircle the planet parallel to its equator.

big bang (429) The theory that the universe began with a violent explosion from which the expanding universe of galaxies eventually formed.

big rip (446) The possible fate of the universe if dark energy increases rapidly and the expansion of space-time pulls galaxies, stars, and ultimately atoms apart.

binary star (274) One of a pair of stars that orbit around their common center of mass.

binding energy (103) The energy needed to pull an electron away from its atom.

biological evolution (458) The combined effect of variation and natural selection resulting in new species arising and existing species adapting to the environment or becoming extinct. Also called Darwinian evolution.

bipolar flow (299) Oppositely directed jets of gas ejected by some protostellar objects.

birth line (302) In the H–R diagram, the line above the main sequence where protostars first become visible.

black dwarf (327) The end state of a white dwarf that has cooled to low temperature.

black hole (357) A mass that has collapsed to such a small volume that its gravity prevents the escape of all radiation; also, the volume of space from which radiation may not escape.

blackbody radiation (107) Radiation emitted by a hypothetical perfect radiator; the spectrum is continuous, and the wavelength of maximum emission depends only on the body's temperature.

blueshift (112) The shortening of the wavelengths of light observed when the source and observer are approaching each other.

Bok globule (300) Small, dark, and dense interstellar cloud only about 1 ly in diameter that contains 10 to 1000 solar masses of gas and dust; thought to be related to star formation.

bottom-up hypothesis (391) The conjecture that the Milky Way Galaxy and other large galaxies formed mostly by collisions and combination of smaller galaxies and star clusters. See also *monolithic collapse* or *top-down hypothesis.*

breccia (177) A rock composed of fragments of earlier rocks bonded together.

bright-line spectrum (110) See *emission spectrum.*

brown dwarf (266, 311) A very cool, low-luminosity star whose mass is not sufficient to ignite nuclear fusion.

butterfly diagram See *Maunder butterfly diagram.*

CAI (241) Calcium–aluminum-rich inclusions found in some meteorites.

calibrate (372) To make observations of reference objects, checks on instrument performance, calculations of units conversions, etc. needed to completely understand measurements of unknown quantities.

Cambrian explosion (461) The sudden appearance of complex life forms at the beginning of the Cambrian period 0.6 to 0.5 billion years ago. Cambrian rocks contain the oldest easily identifiable fossils.

carbonaceous chondrite (240) Stony meteorite that contains both chondrules and volatiles. These may be the least altered remains of the solar nebula still present in the solar system.

carbon–nitrogen–oxygen (CNO) cycle (308) A series of nuclear reactions that use carbon as a catalyst to combine four hydrogen atoms to make one helium atom plus energy; effective in stars more massive than the sun.

Cassegrain focus (88) The optical design of a reflecting telescope in which the secondary mirror reflects light back down the tube through a hole in the center of the objective mirror.

catastrophic hypothesis (142) Explanation for natural processes that depends on dramatic and unlikely events, such as the collision of two stars to produce our solar system.

celestial equator (18) The imaginary line around the sky directly above Earth's equator.

celestial sphere (18) An imaginary sphere of very large radius surrounding Earth to which the planets, stars, sun, and moon seem to be attached.

centaur (247) An outer solar system body with an orbit entirely within the region of the Jovian planets, for example Chiron, that orbits between Saturn and Uranus.

center of mass (67) The balance point of a body or system of bodies.

central bulge (376) The spheroidal cloud of stars at the center of most spiral galaxies including our Milky Way Galaxy.

Cepheid variable star (370) Variable star with a period of 1 to 60 days; the period of variation is related to luminosity.

Chandrasekhar limit (327) The maximum mass of a white dwarf, about 1.4 solar masses; a white dwarf of greater mass cannot support itself and will collapse.

charge-coupled device (CCD) (93) An electronic device consisting of a large array of light-sensitive elements used to record very faint images.

chemical evolution (460) The chemical process that led to the growth of complex molecules on the primitive Earth. This did not involve the reproduction of molecules.

Chicxulub (255) The buried crater associated with the mass extinction event at the end of the age of dinosaurs, named after the town in the coastal region of Mexico's Yucatán peninsula near the center of the crater.

chondrite (241) A stony meteorite that contains chondrules.

chondrule (241) Round, glassy body in some stony meteorites; thought to have solidified very quickly from molten drops of silicate material.

chromatic aberration (78) A distortion found in refracting telescopes because lenses focus different colors at slightly different distances. Images are consequently surrounded by color fringes.

chromosome (457) One of the bodies in a cell that contains the DNA carrying genetic information.

chromosphere (37, 118) Bright gases just above the photosphere of the sun.

circular velocity (66) The velocity required to remain in a circular orbit about a body.

circumpolar constellation (19) Any of the constellations so close to the celestial pole that they never set (or never rise) as seen from a given latitude.

closed orbit (67) An orbit that returns to its starting point; a circular or elliptical orbit. (See *open orbit.*)

closed universe (437) A model universe in which the average density is great enough to stop the expansion and make the universe contract.

CNO cycle (308) See *carbon–nitrogen–oxygen cycle.*

cocoon (297) The cloud of gas and dust around a contracting protostar that conceals it at visible wavelengths.

cold dark matter (442) Invisible matter in the universe composed of heavy, slow-moving particles such as WIMPs.

coma (250) The glowing head of a comet.

comet (148) One of the small, icy bodies that orbit the sun and produce tails of gas and dust when they near the sun.

compact object (209) A star that has collapsed to form a white dwarf, neutron star, or black hole.

comparative planetology (166) The study of planets by comparing the characteristics of different examples.

composite volcano (194) A volcano built up of layers of lava flows and ash falls. These are steep sided and typically associated with subduction zones.

condensation (152) The growth of a particle by addition of material from surrounding gas, one atom or molecule at a time.

condensation sequence (152) The sequence in which different materials condense from the solar nebula at increasing distances from the sun.

constellation (12) One of the stellar patterns identified by name, usually of mythological gods, people, animals, or objects; also, the region of the sky containing that star pattern.

continuous spectrum (110) A spectrum in which there are no absorption or emission lines.

convection (119, 307) Circulation in a fluid driven by heat; hot material rises, and cool material sinks.

convective zone (136) The region inside a star where energy is carried outward as rising hot gas and sinking cool gas.

Copernican Principle (50) The idea that Earth should not be assumed to be in a special location or to be otherwise unique.

corona (37, 118, 192) The faint outer atmosphere of the sun; composed of low-density, very hot, ionized gas. On Venus, round network of fractures and ridges up to 1000 km in diameter, caused by the intrusion of magma below the crust.

coronagraph (121) A telescope designed to photograph the inner corona of the sun.

coronal gas (121) Extremely high-temperature, low-density gas in the interstellar medium.

coronal hole (133) An area of the solar surface that is dark at X-ray wavelengths; thought to be associated with divergent magnetic fields and the source of the solar wind.

coronal mass ejection (CME) (133) Gas trapped in the sun's magnetic field.

cosmic microwave background radiation (430) Radiation from the hot clouds of the big bang explosion. Because of its large redshift, it appears to come from a body whose temperature is only 2.7 K.

cosmic ray (97) A subatomic particle traveling at tremendous velocity that strikes Earth's atmosphere from space.

cosmological constant (Λ) (445) Einstein's constant that represents a repulsion in space to oppose gravity.

cosmological principle (436) The assumption that any observer in any galaxy sees the same general features of the universe.

cosmologist (426) An astronomer or physicist whose research focuses on the overall properties of the universe and its origin.

cosmology (425) The study of the nature, origin, and evolution of the universe.

Coulomb barrier (135) The electrostatic force of repulsion between bodies of like charge; commonly applied to atomic nuclei.

Coulomb force (103) The repulsive force between particles with like electrostatic charge.

critical density (437) The average density of the universe needed to make its curvature flat.

C-type asteroid (245) A type of asteroid common in the outer asteroid belt, with very low reflectivity and grayish color, probably composed of carbonaceous material.

dark age (434) The period of a few hundred million years during which the universe expanded in darkness. Extends from soon after the big bang glow faded into the infrared to the formation of the first stars.

dark energy (445) The energy of empty space that drives the acceleration of the expanding universe.

dark matter (378) Nonluminous material that is detected only by its gravitational influence.

dark nebula (295) A nonluminous cloud of gas and dust visible because it blocks light from more distant stars and nebulae.

dark-line spectrum (110) See *absorption spectrum.*

debris disk (158) A disk of dust found by infrared observations around some stars. The dust is debris from collisions among asteroids, comets, and Kuiper belt objects.

deferent (49) In the Ptolemaic theory, the large circle around Earth along which the center of the epicycle moved.

degenerate matter (320) Extremely high-density matter in which pressure no longer depends on temperature, due to quantum mechanical effects.

density (163) The amount of matter per unit volume in a material; measured in grams per cubic centimeter, for example.

deuterium (134) An isotope of hydrogen in which the nucleus contains a proton and a neutron.

diamond ring effect (38) A momentary phenomenon seen during some total solar eclipses when the ring of the corona and a bright spot of photosphere resemble a large diamond set in a silvery ring.

differential rotation (128) The rotation of a body in which different parts of the body have different periods of rotation; this is true of the sun, the Jovian planets, and the disk of the galaxy.

differentiation (153) The separation of planetary material according to density.

diffraction fringe (80) Blurred fringe surrounding any image caused by the wave properties of light. Because of this, no image detail smaller than the fringe can be seen.

digitize (93) Convert information to numerical form for convenient transfer, storage, and analysis.

direct collapse (155) The hypothetical process by which a Jovian planet might skip the accretion of a solid core, instead forming quickly and directly from the gases of the solar nebula.

disk component (375) All material confined to the plane of the galaxy.

distance indicator (400) Object whose luminosity or diameter is known; used to find the distance to a star cluster or galaxy.

distance scale (401) The combined calibration of distance indicators used by astronomers to find the distances to remote galaxies.

DNA (deoxyribonucleic acid) (456) The long carbon-chain molecule that records information to govern the biological activity of the organism. DNA carries the genetic data passed to offspring.

Doppler effect (109) The change in the wavelength of radiation due to relative radial motion of source and observer.

double-exhaust model (414) The theory that double radio lobes are produced by pairs of jets emitted in opposite directions from the centers of active galaxies.

double-lobed radio galaxy (412) A galaxy that emits radio energy from two regions (lobes) located on opposite sides of the galaxy.

Drake equation (468) A formula for the number of communicative civilizations in our galaxy.

dust tail (250) The tail of a comet formed of dust blown outward by the pressure of sunlight. (See *gas tail.)*

dwarf planet (234) An object that orbits the sun and has pulled itself into a spherical shape but has not cleared its orbital lane of other objects. Pluto is a dwarf planet.

dynamo effect (128) The process by which a rotating, convecting body of conducting matter, such as Earth's core, can generate a magnetic field.

east point (18) The point on the eastern horizon exactly halfway between the north point and the south point; exactly east.

eccentric (57) (noun) An off-center circular path. (Note that the adjective "eccentric" refers instead to an ellipse that is not a perfect circle.)

eccentricity, e (57) A measure of the flattening of an ellipse. An ellipse of $e = 0$ is circular. The closer to 1 that e becomes, the more flattened the ellipse.

eclipse season (41) That period when the sun is near a node of the moon's orbit and eclipses are possible.

eclipsing binary (278) A binary star system in which the stars eclipse each other.

ecliptic (25) The apparent path of the sun around the sky.

ejecta (178) Pulverized rock scattered by meteorite impacts on a planetary surface.

electromagnetic radiation (75) Changing electric and magnetic fields that travel through space and transfer energy from one place to another—for example, light, radio waves, and the like.

electron (102) Low-mass atomic particle carrying a negative charge.

ellipse (56) A closed curve enclosing two points (foci) such that the total distance from one focus to any point on the curve back to the other focus equals a constant.

elliptical galaxy (398) A galaxy that is round or elliptical in outline; it contains little gas and dust, no disk or spiral arms, and few hot, bright stars.

emission (bright-line) spectrum (110) A spectrum containing emission lines.

emission line (110) A bright line in a spectrum caused by the emission of photons from atoms.

emission nebula (294) A cloud of glowing gas excited by ultraviolet radiation from hot stars.

energy level (105) One of a number of states an electron may occupy in an atom, depending on its binding energy.

energy transport (307) The law of energy transport states that energy must flow from hot regions to cool regions by conduction, convection, or radiation.

enzyme (456) Special protein that controls processes in an organism.

epicycle (49) The small circle followed by a planet in the Ptolemaic theory. The center of the epicycle follows a larger circle (deferent) around Earth.

equant (49) The point off-center in the deferent from which the center of the epicycle appears to move uniformly.

equatorial mount (89) A telescope mounting that allows motion parallel to and perpendicular to the celestial equator.

escape velocity (V_e) (67) The initial velocity an object needs to escape from the surface of a celestial body.

evening star (25) Any planet visible in the sky just after sunset.

event horizon (358) The boundary of the region of a black hole from which no radiation may escape. No event that occurs within the event horizon is visible to a distant observer.

evolutionary hypothesis (142) Explanation for natural events that involves gradual changes as opposed to sudden catastrophic changes—for example, the formation of the planets in the gas cloud around the forming sun.

excited atom (105) An atom in which an electron has moved from a lower to a higher orbit.

expanding universe (428) The idea, supported by observed redshifts of galaxies, that space is stretching, carrying galaxies and galaxy clusters away from each other.

extrasolar planet (159) A planet orbiting a star other than the sun.

extremophile (462) An organism that can survive in an extreme environment, for example, very low or high temperatures, high acidity, extreme dryness, and so on.

eyepiece (78) A short-focal-length lens used to enlarge the image in a telescope; the lens nearest the eye.

false-color image. See *representational-color image.*

field (93) A way of explaining action at a distance; a particle produces a field of influence (gravitational, electric, or magnetic) to which another particle in the field responds.

field of view (2) The area visible in an image; usually given as the diameter of the region.

filament (132, 446) (1) On the sun, a prominence seen silhouetted against the solar surface. (2) A linear region containing many galaxies and galaxy clusters, part of the large-scale structure of the universe.

filtergram (121) An image (usually of the sun) taken in the light of a specific region of the spectrum—for example, an H-alpha filtergram.

flare (133) A violent eruption on the sun's surface.

flat universe (437) A model of the universe in which space-time is not curved.

flatness problem (442) In cosmology, the circumstance that the early universe must have contained almost exactly the right amount of matter to close space-time (to make space-time flat).

flocculent (382) Woolly, fluffy; used to refer to certain galaxies that have a woolly appearance.

flux (15, 262) A measure of the flow of energy onto or through a surface. Usually applied to light.

focal length (78) The distance from a lens to the point where it focuses parallel rays of light.

folded mountain range (172) A long range of mountains formed by the compression of a planet's crust—for example, the Andes on Earth.

forward scattering (216) The optical property of finely divided particles to preferentially direct light in the original direction of the light's travel.

free-fall collapse (296) The early contraction of a gas cloud to form a star during which internal pressure is too low to resist contraction.

frequency (75) The number of times a given event occurs in a given time; for a wave, the number of cycles that pass the observer in 1 second.

galactic cannibalism (410) The theory that large galaxies absorb smaller galaxies.

galactic corona (377) The low-density extension of the halo of a galaxy; now suspected to extend many times the visible diameter of the galaxy.

galaxy (5) A very large collection of gas, dust, and stars orbiting a common center of mass. The sun and Earth are located in the Milky Way Galaxy.

Galilean moons or satellites (60, 213) The four largest satellites of Jupiter, named after their discoverer, Galileo.

gamma ray (76) Electromagnetic wave with extremely short wavelength, high frequency, and large photon energy.

gamma-ray burst (363) A sudden burst of gamma rays thought to be associated with neutron stars and black holes.

gas tail (250) The tail of a comet produced by gas blown outward by the solar wind. (See *dust tail.*)

gene (457) A unit of DNA containing genetic information that influences a particular inherited trait.

general theory of relativity (98) Einstein's more sophisticated theory of space and time, which describes gravity as a curvature of space-time.

geocentric universe (48) A model universe with Earth at the center, such as the Ptolemaic universe.

geosynchronous satellite (66) An Earth satellite in an eastward orbit whose period is 24 hours. A satellite in such an orbit remains above the same spot on Earth's surface.

giant (272) Large, cool, highly luminous star in the upper right of the H–R diagram, typically 10 to 100 times the diameter of the sun.

giant molecular cloud (296) Very large, cool cloud of dense gas in which stars form.

global warming (171) The gradual increase in the surface temperature of Earth caused by human modifications to Earth's atmosphere.

globular cluster (324) A star cluster containing 50,000 to 1 million stars in a sphere about 75 ly in diameter; generally old, metal-poor, and found in the spherical component of the galaxy.

grand design (382) Galaxy with a high-contrast, simple, two-arm spiral pattern.

grand unified theory (GUT) (443) Theory that attempts to unify (describe in a similar way) the electromagnetic, weak, and strong forces of nature.

granulation (119) The fine structure visible on the solar surface caused by rising currents of hot gas and sinking currents of cool gas below the surface.

grating (94) A piece of material in which numerous microscopic parallel lines are scribed; light encountering a grating is dispersed to form a spectrum.

gravitational collapse (153) The stage in the formation of a massive planet when it grows massive enough to begin capturing gas directly from the nebula around it.

gravitational lensing (405) The effect of the focusing of light from a distant galaxy or quasar by an intervening galaxy to produce multiple images of the distant body.

gravitational radiation (352) As predicted by general relativity, expanding waves in a gravitational field that transport energy through space.

gravitational redshift (359) The lengthening of the wavelength of a photon due to its escape from a gravitational field.

greenhouse effect (171) The process by which a carbon dioxide atmosphere traps heat and raises the temperature of a planetary surface.

grooved terrain (213) Region of the surface of Ganymede consisting of bright, parallel grooves.

ground state (105) The lowest permitted electron orbit in an atom.

habitable zone (465) The region around a star within which an orbiting planet can have surface temperatures allowing liquid water.

half-life (149) The time required for half of the atoms in a radioactive sample to decay.

halo (376) The spherical region of a spiral galaxy containing a thin scattering of stars, star clusters, and small amounts of gas.

heat (106) Energy flowing from a warm body to a cool body by the agitation of particles such as atoms or molecules.

heat of formation (153) In planetology, the heat released by the infall of matter during the formation of a planetary body.

heavy bombardment (157) The period of intense meteorite impacts early in the formation of the planets, when the solar system was filled with debris.

heliocentric universe (50) A model of the universe with the sun at the center, such as the Copernican universe.

helioseismology (123) The study of the interior of the sun by the analysis of its modes of vibration.

helium flash (321) The explosive ignition of helium burning that takes place in some giant stars.

Herbig–Haro object (299) A small nebula associated with star formation that varies irregularly in brightness.

Hertzsprung–Russell (H–R) diagram (269) A plot of the intrinsic brightness versus the surface temperature of stars; it separates the effects of temperature and surface area on stellar luminosity; commonly absolute magnitude versus spectral type but also luminosity versus surface temperature or color.

H II region (294) A region of ionized hydrogen around a hot star.

homogeneous (436) The property of being uniform. In cosmology, the characteristic of the universe in which, on the large scale, matter is uniformly spread through the universe.

horizon (16) The line that marks the apparent intersection of Earth and the sky.

horizon problem (442) In cosmology, the circumstance that the primordial background radiation seems much more isotropic than could be explained by the standard big bang theory.

horizontal branch (325) In the H–R diagram of a globular cluster, the sequence of stars extending from the red giants toward the blue side of the diagram; includes RR Lyrae stars.

horoscope (28) A chart showing the positions of the sun, moon, planets, and constellations at the time of a person's birth; used in astrology to attempt to read character or foretell the future.

hot dark matter (442) Invisible matter in the universe composed of low-mass, high-velocity particles such as neutrinos.

hot Jupiter (161) A massive and presumably Jovian planet that orbits close to its star and consequently has a high temperature.

hot spot (414) In radio astronomy, a bright spot in a radio lobe.

H–R diagram (269) See *Hertzsprung–Russell diagram.*

Hubble constant (*H*) (402) A measure of the rate of expansion of the universe; the average value of velocity of recession divided by distance; about 70 km/s/megaparsec.

Hubble law (402) The linear relation between the distance to a galaxy and its radial velocity.

Hubble time (429) An upper limit on the age of the universe derived from the Hubble constant.

hydrostatic equilibrium (306) The balance between the weight of the material pressing downward on a layer in a star and the pressure in that layer.

hypernova (364) The explosion produced as a very massive star collapses into a black hole; thought to be responsible for at least some gamma-ray bursts.

hypothesis (58) A conjecture, subject to further tests, that accounts for a set of facts.

ice line (152) In the solar nebula, the boundary beyond which water vapor and other compounds could form ice particles.

inflationary universe (443) A version of the big bang theory that includes a rapid expansion when the universe was very young.

infrared (IR) radiation (76) Electromagnetic radiation with wavelengths intermediate between visible light and radio waves.

inner Lagrange point (331) The point of gravitational equilibrium between two orbiting stars through which matter can flow from one star to the other.

instability strip (370) The region of the H–R diagram in which stars are unstable to pulsation; a star passing through this strip becomes a variable star.

interferometry (95) The observing technique in which separated telescopes combine to produce a virtual telescope with the resolution of a much-larger-diameter telescope.

International Astronomical Union (IAU) (223) An international society of astronomers that, among other activities, decides definitions and naming conventions for celestial objects and surface features. The IAU defined the constellation boundaries in 1930 and reclassified Pluto as a dwarf planet in 2006.

interstellar absorption line (290) One of the dark lines in some stellar spectra that are formed by interstellar gas.

interstellar dust (290) Microscopic solid grains in the interstellar medium.

interstellar extinction (289) The dimming of starlight by gas and dust in the interstellar medium.

interstellar medium (ISM) (289) The gas and dust distributed between the stars.

interstellar reddening (289) The process in which dust scatters blue light out of starlight and makes the stars look redder.

intrinsic brightness (262) The true brightness of an object independent of its distance. Also referred to as luminosity.

inverse square relation (63) The rule that the strength of an effect (such as gravity) decreases in proportion as the distance squared increases.

ion (103) An atom that has lost or gained one or more electrons.

ionization (103) The process in which atoms lose or gain electrons.

iron meteorite (239) A meteorite composed mainly of iron–nickel alloy.

irregular galaxy (399) A galaxy with a chaotic appearance, large clouds of gas and dust, and both population I and population II stars, but without spiral arms.

irregular satellite (211) A moon with an orbit that has large eccentricity and/or high inclination to the equator of its parent planet and/or is retrograde. Irregular moons are thought to have been captured.

isotopes (103) Atoms that have the same number of protons but a different number of neutrons.

isotropic (436) The condition of being uniform in all directions. In cosmology, the characteristic of the universe by which, in its general properties, it looks the same in every direction.

Jovian planet (146) Jupiter-like planet with large diameter and low density.

Jovian problem (155) The puzzle that protoplanetary disks around young stars don't seem to survive long enough to form Jovian planets by condensation, accretion, and gravitational collapse, yet Jovian-mass extrasolar planets are common. See also *extrasolar planet.*

Kelvin temperature scale (107) The temperature, in Celsius (centigrade) degrees, measured above absolute zero.

Keplerian motion (377) Orbital motion in accord with Kepler's laws of planetary motion.

kiloparsec (kpc) (369) A unit of distance equal to 1000 pc, or 3260 ly.

Kirchhoff's laws (110) A set of laws that describe the absorption and emission of light by matter.

Kuiper belt (148) The collection of icy planetesimals that orbit in a region from just beyond Neptune out to about 50 AU.

Kuiper belt object (148) An object in the Kuiper belt, a region beyond Neptune's orbit containing planetesimals remaining from the formation of the solar system. Pluto is one of the largest Kuiper belt objects.

L dwarf (266) A type of star that is even cooler than the M stars.

Lagrange point (331) Point of stability in the orbital plane of a binary star system, planet, or moon. One is located 60° ahead and one 60° behind the orbiting bodies; another is located between the orbiting bodies.

large-impact hypothesis (180) The hypothesis that the moon formed from debris ejected during a collision between Earth and a large planetesimal.

large-scale structure (446) The distribution of galaxy clusters and superclusters in walls and filaments surrounding voids mostly empty of galaxies.

laser guide star (95) An artificial star image produced by a laser pointing up into Earth's atmosphere, used as a reference in adaptive optics systems. See also *adaptive optics.*

late heavy bombardment (181) The surge in cratering impacts in the solar system that occurred about 3.8 billion years ago.

light curve (279) A graph of brightness versus time commonly used in analyzing variable stars and eclipsing binaries.

light pollution (83) The illumination of the night sky by waste light from cities and outdoor lighting, which prevents the observation of faint objects.

light-gathering power (79) The ability of a telescope to collect light; proportional to the area of the telescope objective lens or mirror.

lighthouse model (347) The explanation of a pulsar as a spinning neutron star sweeping beams of radio radiation around the sky.

light-year (ly) (4) The distance light travels in one year.

limb (144, 450) The edge of the apparent disk of a body, as in "the limb of the moon."

liquid metallic hydrogen (211) A form of hydrogen under high pressure that is a good electrical conductor.

lobate scarp (189) A curved cliff such as those found on Mercury.

look-back time (402) The amount by which you look into the past when you look at a distant galaxy; a time equal to the distance to the galaxy in light-years.

luminosity (*L*) (264) The total amount of energy a star radiates in 1 second.

luminosity class (272) A category of stars of similar luminosity; determined by the widths of lines in their spectra.

lunar eclipse (33) The darkening of the moon when it moves through Earth's shadow.

Lyman series (111) Spectral lines in the ultraviolet spectrum of hydrogen produced by transitions whose lowest orbit is the ground state.

magnetar (364) A class of neutron stars that have exceedingly strong magnetic fields; thought to be responsible for soft gamma-ray repeaters.

magnetic carpet (122) The widely distributed, low-level magnetic field extending up through the sun's visible surface.

magnetosphere (211) The volume of space around a planet within which the motion of charged particles is dominated by the planetary magnetic field rather than the solar wind.

magnifying power (82) The ability of a telescope to make an image larger.

magnitude scale (14) The astronomical brightness scale; the larger the number, the fainter the star.

magnitude–distance formula (263) The mathematical formula that relates the apparent magnitude and absolute magnitude of a star to its distance.

mantle (166) The layer of dense rock and metal oxides that lies between the molten core and Earth's surface; also, similar layers in other planets.

mare (mä´rā) (177) (plural: **maria**) One of the lunar lowlands filled by successive flows of dark lava; from the Latin word for sea.

mass (63) A measure of the amount of matter making up an object.

mass–luminosity relation (281) The more massive a star is, the more luminous it is.

Maunder butterfly diagram (126) A graph showing the latitude of sunspots versus time; first plotted by W. W. Maunder in 1904.

Maunder minimum (127) A period of less numerous sunspots and other solar activity from 1645 to 1715.

megaparsec (Mpc) (400) A unit of distance equal to 1 million pc.

metal (388) In astronomical usage, any atom heavier than helium.

meteor (342) A small bit of matter heated by friction to incandescent vapor as it falls into Earth's atmosphere.

meteor shower (242) An event lasting for hours or days in which the number of meteors entering Earth's atmosphere suddenly increases. The meteors in a shower have a common origin and are traveling through space on nearly parallel paths.

meteorite (149) A meteor that has survived its passage through the atmosphere and strikes the ground.

meteoroid (149) A meteor in space before it enters Earth's atmosphere.

microlensing (160) Brightening of a background star due to focusing of its light by the gravity of a foreground extrasolar planet, allowing the planet to be detected and some of its characteristics measured.

micrometeorite (179) Meteorite of microscopic size.

microwave (76) Electromagnetic wave with wavelength, frequency, and photon energy intermediate between infrared and radio waves.

mid-ocean rise (172) One of the undersea mountain ranges that push up from the seafloor in the center of the oceans.

Milankovitch hypothesis (28) The hypothesis that small changes in Earth's orbital and rotational motions cause the ice ages.

Milky Way (5) The hazy band of light that circles the sky, produced by the combined light of billions of stars in our Milky Way Galaxy.

Milky Way Galaxy (5) The spiral galaxy containing the sun; visible at night as the Milky Way.

Miller experiment (459) An experiment that reproduced the conditions under which life began on Earth and amino acids and other organic compounds were manufactured.

millisecond pulsar (354) A pulsar with a period of approximately a millisecond, a thousandth of a second.

molecular cloud (290) An interstellar gas cloud that is dense enough for the formation of molecules; discovered and studied through the radio emissions of such molecules.

molecule (103) Two or more atoms bonded together.

monolithic collapse hypothesis (390) The hypothesis that the Milky Way Galaxy formed by gravitational collapse of a single large spinning cloud of gas. This hypothesis is now considered inadequate to explain many observed characteristics of the galaxy.

morning star (25) Any planet visible in the sky just before sunrise.

M-type asteroid (245) A type of asteroid with relatively high reflectivity and grayish color, probably composed primarily of metal.

multicellular (461) An organism composed of many cells.

multiringed basin (179) Very large impact basin in which there are concentric rings of mountains.

mutation (458) Offspring born with altered DNA.

nanometer (nm) (76) A unit of length equal to 10^{-9} m.

natural law (58) A conjecture about how nature works in which scientists have overwhelming confidence.

natural selection (458) The process by which the best traits are passed on, allowing the most able to survive.

neap tide (70) Ocean tide of low amplitude occurring at first- and third-quarter moon.

Near-Earth Object (NEO) (246) An asteroid or comet in an orbit that passes near or intersects Earth's orbit, that could potentially collide with Earth.

nebula (292) A cloud of gas and dust in space.

neutrino (134) A neutral, massless atomic particle that travels at or nearly at the speed of light.

neutron (102) An atomic particle with no charge and about the same mass as a proton.

Newtonian focus (88) The focal arrangement of a reflecting telescope in which a diagonal mirror reflects light out the side of the telescope tube for easier access.

node (40) A point where an object's orbit passes through the plane of Earth's orbit.

nonbaryonic matter (440) In cosmology, a suspected component of the dark matter composed of matter that does not contain protons and neutrons.

nova (318) From the Latin "new," a sudden brightening of a star, making it appear as a "new" star in the sky; thought to be associated with eruptions on white dwarfs in binary systems.

nuclear fission (131) Reaction that splits nuclei into less massive fragments.

nuclear fusion (131) Reaction that joins the nuclei of atoms to form more massive nuclei.

nucleosynthesis (388) The production of elements heavier than helium by the fusion of atomic nuclei in stars and during supernovae explosions.

nucleus (of an atom) (102) The central core of an atom containing protons and neutrons; carries a net positive charge.

OB association (303) A loosely bound cluster of young stars having spectral types O and B, indicating a region of relatively recent star formation.

oblateness (211) The flattening of a spherical body, usually caused by rotation.

observable universe (428) The part of the universe that is visible from Earth's location in space and time.

occultation (226) The passage of a larger body in front of a smaller body.

Olbers's paradox (426) The conflict between observation and theory as to why the night sky should or should not be dark.

opacity (307) The resistance of a gas to the passage of radiation.

open cluster (324) A cluster of 10 to 10,000 stars with an open, transparent appearance and stars not tightly grouped; usually relatively young and located in the disk of the galaxy.

open orbit (67) An orbit that does not return to its starting point; an escape orbit. (See *closed orbit.*)

open universe (437) A model universe in which the average density is less than the critical density needed to halt the expansion.

optical telescope (79) A telescope that gathers and focuses visible light. See also *radio telescope.*

outflow channel (201) Geological feature on Mars that appears to have been caused by sudden flooding.

outgassing (154) The release of gases from a planet's interior.

ovoid (223) Geological feature on Uranus's moon Miranda thought to be produced by circulation in the solid icy mantle and crust.

ozone layer (171) In Earth's atmosphere, a layer of oxygen ions (O_3) lying 15 to 30 km high that protects the surface by absorbing ultraviolet radiation.

paradigm (53) A commonly accepted set of scientific ideas and assumptions.

parallax (48, 169) The apparent change in the position of an object due to a change in the location of the observer. Astronomical parallax is measured in seconds of arc.

Paschen series (111) Spectral lines in the infrared spectrum of hydrogen produced by transitions whose lowest orbit is the third.

penumbra (33) The portion of a shadow that is only partially shaded.

perigee (38) The orbital point of closest approach to Earth.

perihelion (27) The orbital point of closest approach to the sun.

period–luminosity relation (371) The relation between period of pulsation and intrinsic brightness among Cepheid variable stars.

permitted orbit (104) One of the energy levels in an atom that an electron may occupy.

photographic plate (93) An old-fashioned means of recording astronomical images and photometric information on a photographic emulsion coating a glass plate. See also *array detector* and *charge-coupled device.*

photometer (93) An instrument attached to a telescope for the purpose of precisely measuring the brightness of stars or other objects at one or more wavelengths.

photon (76) A quantum of electromagnetic energy; carries an amount of energy that depends inversely on its wavelength.

photosphere (37, 118) The bright visible surface of the sun.

planetary nebula (326) An expanding shell of gas ejected from a star during the latter stages of its evolution.

planetesimal (152) One of the small bodies that formed from the solar nebula and eventually grew into protoplanets.

plastic (170) A material with the properties of a solid but capable of flowing under pressure.

plate tectonics (172) The constant destruction and renewal of Earth's surface by the motion of sections of crust.

plutino (234) One of the icy Kuiper belt objects that, like Pluto, are caught in a 3:2 orbital resonance with Neptune.

polar axis (89) In an equatorial telescope mounting, the axis that is parallel to Earth's axis of rotation.

polarity (128) Orientation and strength of a magnetic field's manifestation as north and south poles. Also applies to an electrical field's manifestation as positive and negative charges.

poor cluster (406) An irregularly shaped cluster that contains fewer than 1000 galaxies, many spiral, and no giant ellipticals.

Population I star (388) Star rich in atoms heavier than helium; nearly always a relatively young star found in the disk of the galaxy.

Population II star (388) Star poor in atoms heavier than helium; nearly always a relatively old star found in the halo, globular clusters, or the nuclear bulge.

positron (134) The antiparticle of the electron.

potential energy (89) The energy a body has by virtue of its position. A weight on a high shelf has more potential energy than a weight on a low shelf.

precession (20) The slow change in the direction of Earth's axis of rotation; one cycle takes nearly 26,000 years.

pressure *(P)* wave (169) In geophysics, a mechanical wave of compression and rarefaction that travels through Earth's interior.

pressure–temperature thermostat (309) The dependence of gas pressure on gas temperature, which results in stability and regulation of energy production in the cores of normal stars.

primary lens or mirror (78) The main optical element in an astronomical telescope. The large lens at the top of the telescope tube or the large mirror at the bottom.

prime focus (88) The point at which the objective mirror forms an image in a reflecting telescope.

primordial soup (459) The rich solution of organic molecules in Earth's first oceans.

prograde (211) Rotation or revolution in the direction in common with most such motions in the solar system. See also *retrograde.*

prominence (37, 132) Eruption on the solar surface; visible during total solar eclipses.

proper motion (372) The rate at which a star moves across the sky; measured in seconds of arc per year.

protein (456) Complex molecule composed of amino acid units.

proton (102) A positively charged atomic particle contained in the nucleus of an atom; the nucleus of a hydrogen atom.

proton–proton chain (134) A series of three nuclear reactions that build a helium atom by adding together protons; the main energy source in the sun.

protoplanet (153) Massive object resulting from the coalescence of planetesimals in the solar nebula and destined to become a planet.

protostar (297) A collapsing cloud of gas and dust destined to become a star.

protostellar disk (298) A gas cloud around a forming star flattened by its rotation.

pulsar (347) A source of short, precisely timed radio bursts; thought to be a spinning neutron star.

pulsar wind (350) The flow of high-energy particles that carries most of the energy away from a spinning neutron star.

quantum mechanics (104) The study of the behavior of atoms and atomic particles.

quasar (quasi-stellar object, or QSO) (414) Small, powerful source of energy thought to be the active core of a very distant galaxy.

quintessence (439) The proposed energy of empty space that causes the acceleration of the expanding universe.

radial velocity (V_r) (114) That component of an object's velocity directed away from or toward Earth.

radiant (242) The point in the sky from which meteors in a shower seem to come.

radiation pressure (148, 307) The force exerted on the surface of a body by its absorption of light. Small particles floating in the solar system can be blown outward by the pressure of the sunlight.

radiative zone (136) The region inside a star where energy is carried outward as photons.

radio galaxy (409) A galaxy that is a strong source of radio signals.

radio telescope (79) A telescope that gathers and focuses electromagnetic energy with microwave and radio wavelengths. See also *optical telescope.*

radio wave (76) Electromagnetic wave with extremely long wavelength, low frequency, and small photon energy.

ray (178) Ejecta from a meteorite impact, forming white streamers radiating from some lunar craters.

recombination (433) The stage within a million years of the big bang when the gas became transparent to radiation.

reconnection event (133) The process in the sun's atmosphere by which opposing magnetic fields combine and release energy to power solar flares.

red dwarf (272) Cool, low-mass star on the lower main sequence.

redshift (112) The lengthening of the wavelengths of light seen when the source and observer are receding from each other.

reflecting telescope (78) A telescope that uses a concave mirror to focus light into an image.

reflection nebula (294) A nebula produced by starlight reflecting off dust particles in the interstellar medium.

refracting telescope (78) A telescope that forms images by bending (refracting) light with a lens.

regular satellite (211) A moon with an orbit that has small eccentricity, low inclination to the equator of its parent planet, and is prograde. Regular moons are thought to have formed with their respective planets rather than having been captured.

re-ionization (434) The stage in the early history of the universe when ultraviolet photons from the first stars ionized the gas filling space.

representational-color image (93) A representation of graphical data in which the colors are altered or added to reveal details. Also sometimes called a *false-color image.*

resolving power (80) The ability of a telescope to reveal fine detail; depends on the diameter of the telescope objective.

retrograde motion (48) The apparent backward (westward) motion of planets as seen against the background of stars.

revolution (24) The motion of an object in a closed path about a point outside its volume; Earth revolves around the sun.

rich cluster (406) A cluster containing more than 1000 galaxies, mostly elliptical, scattered over a volume about 3 Mpc in diameter.

rift valley (172) A long, straight, deep valley produced by the separation of crustal plates.

ring galaxy (411) A galaxy that resembles a ring around a bright nucleus; thought to be the result of a head-on collision of two galaxies.

RNA (ribonucleic acid) (457) A long carbon-chain molecule that uses the information stored in DNA to manufacture complex molecules necessary to the organism.

Roche limit (217) The minimum distance between a planet and a satellite that holds itself together by its own gravity. If a satellite's orbit brings it within its planet's Roche limit, tidal forces will pull the satellite apart.

Roche lobe (331) In a system with two bodies orbiting each other, the volume of space dominated by the gravitation of one of the bodies.

Roche surface (331) In a system with two bodies orbiting each other, the outer boundary of the volume of space dominated by the gravitation of one of the bodies.

rotation (24) The turning of a body about an axis that passes through its volume; Earth rotates on its axis.

rotation curve (337, 403) A graph of orbital velocity versus radius in the disk of a galaxy.

rotation curve method (403) The procedure for finding the mass of a galaxy from its rotation curve.

RR Lyrae variable star (371) Variable star with a period of 12 to 24 hours; common in some globular clusters.

Sagittarius A* (384) The powerful radio source located at the core of the Milky Way Galaxy.

Saros cycle (42) An 18-year 11 ⅓-day period after which the pattern of lunar and solar eclipses repeats.

Schmidt-Cassegrain focus (88) The optical design of a reflecting telescope in which a thin correcting lens is placed at the top of a Cassegrain telescope.

Schwarzschild radius (R_S) (358) The radius of the event horizon around a black hole.

scientific argument (31) An honest, logical discussion of observations and theories intended to reach a valid conclusion.

scientific method (7) The reasoning style by which scientists test theories against evidence to understand how nature works.

scientific notation (3) The system of recording very large or very small numbers by using powers of 10.

secondary atmosphere (171) The gases outgassed from a planet's interior; rich in carbon dioxide.

secondary crater (178) A crater formed by the impact of debris ejected from a larger crater.

secondary mirror (88) In a reflecting telescope, the mirror that reflects the light to a point of easy observation.

seeing (80) Atmospheric conditions on a given night. When the atmosphere is unsteady, producing blurred images, the seeing is said to be poor.

seismic wave (169) A mechanical vibration that travels through Earth; usually caused by an earthquake.

seismograph (169) An instrument that records seismic waves.

selection effect (240) An influence on the probability that certain phenomena will be detected or selected, which can alter the outcome of a survey.

self-sustaining star formation (382) The process by which the birth of stars compresses the surrounding gas clouds and triggers the formation of more stars; proposed to explain spiral arms.

semimajor axis *(a)* (57) Half of the longest axis of an ellipse.

SETI (467) Search for Extra-Terrestrial Intelligence.

Seyfert galaxy (409) An otherwise normal spiral galaxy with an unusually bright, small core that fluctuates in brightness; thought to indicate the core is erupting.

shear *(S)* wave (170) A mechanical wave that travels through Earth's interior by the vibration of particles perpendicular to the direction of wave travel.

shepherd satellite (225) A satellite that, by its gravitational field, confines particles to a planetary ring.

shield volcano (194) Wide, low-profile volcanic cone produced by highly liquid lava.

shock wave (293) A sudden change in pressure that travels as an intense sound wave.

sidereal drive (89) The motor and gears on a telescope that turn it westward to keep it pointed at a star.

sidereal period (35) The period of rotation or revolution of an astronomical body relative to the stars.

singularity (357) The object of zero radius into which the matter in a black hole is thought to fall.

small-angle formula (38) The mathematical formula that relates an object's linear diameter and distance to its angular diameter.

solar constant (130) A measure of the energy output of the sun; the total solar energy striking 1 m^2 just above Earth's atmosphere in 1 second.

solar eclipse (37) The event that occurs when the moon passes directly between Earth and the sun, blocking your view of the sun.

solar nebula theory (142) The proposal that the planets formed from the same cloud of gas and dust that formed the sun.

solar wind (122) Rapidly moving atoms and ions that escape from the solar corona and blow outward through the solar system.

south celestial pole (18) The point of the celestial sphere directly above Earth's South Pole.

south point (18) The point on the horizon directly above the south celestial pole; exactly south.

spectral line (94) A dark or bright line that crosses a spectrum at a specific wavelength.

spectral sequence (266) The arrangement of spectral classes (O, B, A, F, G, K, M) ranging from hot to cool.

spectral type or class (266) A star's position in the temperature classification system O, B, A F, G, K, and M. Based on the appearance of the star's spectrum.

spectrograph (94) A device that separates light by wavelength to produce a spectrum.

spectroscopic binary (276) A star system in which the stars are too close together to be visible separately. You see a single point of light, and only by taking a spectrum can you determine that there are two stars.

spectroscopic parallax (274) The method of determining a star's distance by comparing its apparent magnitude with its absolute magnitude, as estimated from its spectrum.

spectrum (76) An arrangement of electromagnetic radiation in order of wavelength or frequency.

spherical component (376) The part of the galaxy including all matter in a spherical distribution around the center (the halo and nuclear bulge).

spicule (121) Small, flamelike projection in the chromosphere of the sun.

spiral arm (6) Long, spiral pattern of bright stars, star clusters, gas, and dust that extends from the center to the edge of the disk of spiral galaxies.

spiral density wave theory (381) The conjecture that spiral arms in disk galaxies are caused by a pressure wave that rotates slowly around the galaxy, triggering star formation by compressing interstellar gas clouds.

spiral galaxy (398) A galaxy with an obvious disk component containing gas; dust; hot, bright stars; and spiral arms.

spiral tracer (379) Object used to map the spiral arms, for example O and B associations, open clusters, clouds of ionized hydrogen, and some types of variable stars.

sporadic meteor (242) A meteor not part of a meteor shower.

spring tide (70) Ocean tide of high amplitude that occurs at full and new moon.

standard candle (400) Object of known brightness that astronomers use to find distance—for example, Cepheid variable stars and supernovae.

starburst galaxy (408) A bright blue galaxy in which many new stars are forming, thought to be caused by collisions between galaxies.

static (426) Unchanging in overall properties; opposite of evolving.

Stefan-Boltzmann law (108) The mathematical relation between the temperature of a blackbody (an ideal radiator) and the amount of energy emitted per second from 1 square meter of its surface.

stellar model (309) A table of numbers representing the conditions in various layers within a star.

stellar parallax *(p)* (261) A measure of stellar distance. (*See parallax.*)

stony meteorite (239) A meteorite composed of silicate (rocky) material.

stony-iron meteorite (239) A meteorite that is a mixture of stone and iron.

stromatolite (459) A layered fossil formation caused by ancient mats of algae or bacteria that build up mineral deposits season after season.

strong force (131) One of the four forces of nature; the strong force binds protons and neutrons together in atomic nuclei.

S-type asteroid (245) A type of asteroid common in the inner asteroid belt, with relatively high reflectivity and reddish color, probably composed of rocky material.

subduction zone (172) A region of a planetary crust where a tectonic plate slides downward.

summer solstice (26) The point on the celestial sphere where the sun is at its most northerly point; also, the time when the sun passes this point, about June 22, and summer begins in the Northern Hemisphere.

sunspot (118) Relatively dark spot on the sun that contains intense magnetic fields.

supercluster (446) A cluster of galaxy clusters.

supergiant (272) Exceptionally luminous star, 10 to 1000 times the sun's diameter.

supergranule (119) A large granule on the sun's surface including many smaller granules.

supernova (type I) (337) The violent explosion of a star in which the spectrum contains no hydrogen lines.

supernova (type Ia) (337) The explosion of a star caused by the collapse of a white dwarf that has gained mass from its binary companion and exceeds the Chandrasekhar limit.

supernova (type Ib) (337) The explosion of a massive star that develops an iron core and collapses after it has lost its outer layers of hydrogen.

supernova (type II) (337) The explosion of a massive star that develops an iron core and collapses.

supernova remnant (340) The expanding shell of gas marking the site of a supernova explosion.

synchrotron radiation (337) Radiation emitted when high-speed electrons move through a magnetic field.

T association (303) A large, loosely bound group of T Tauri stars.

T dwarf (266) A very low-mass star at the bottom end of the main sequence with a cool surface and a low luminosity.

T Tauri star (304) Young star surrounded by gas and dust contracting toward the main sequence.

temperature (106) A measure of the velocity of random motions among the atoms or molecules in a material.

terminator (177) The dividing line between daylight and darkness on a planet or moon.

Terrestrial planet (146) Earth-like planet—small, dense, rocky.

theory (58) A system of assumptions and principles applicable to a wide range of phenomena that have been repeatedly verified.

thermal energy (106) The energy stored in an object as agitation among its atoms and molecules.

tidal heating (216) The heating of a planet or satellite because of friction caused by tides.

time dilation (359) The slowing of moving clocks or clocks in strong gravitational fields.

top-down hypothesis (390) See *monolithic collapse hypothesis.*

totality (33) The period during a solar eclipse when the sun's photosphere is completely hidden by the moon, or the period during a lunar eclipse when the moon is completely inside the umbra of Earth's shadow.

transit (160) The passage of an extrasolar planet across the disk of its parent star as observed from Earth, partially blocking the light from the star and allowing detection and study of the planet.

transition (111) The movement of an electron from one atomic orbit to another.

triple alpha process (320) The nuclear fusion process that combines three helium nuclei (alpha particles) to make one carbon nucleus.

Trojan asteroid (247) Small, rocky body caught in Jupiter's orbit at the Lagrange points, 60° ahead of and behind the planet.

turnoff point (324) The point in an H–R diagram where a cluster's stars turn off the main sequence and move toward the red giant region, revealing the approximate age of the cluster.

21-cm radiation (291) Radio emission produced by cold, low-density hydrogen in interstellar space.

ultraviolet radiation (76) Electromagnetic radiation with wavelengths shorter than visible light but longer than X-rays.

umbra (33) The region of a shadow that is totally shaded.

uncompressed density (151) The density a planet would have if its gravity did not compress it.

unified model (418) The attempt to explain the different kinds of active galaxies and quasars by a single model.

uniform circular motion (48) The classical belief that the perfect heavens could move only by the combination of constant motion along circular orbits.

valley networks (201) Dry drainage channels resembling streambeds found on Mars.

vernal equinox (26) The place on the celestial sphere where the sun crosses the celestial equator moving northward; also, the time of year when the sun crosses this point, about March 21, and spring begins in the Northern Hemisphere.

vesicular (177) A porous basalt rock formed by solidified lava with trapped bubbles.

void (446) A region containing relatively few galaxies, part of the large-scale structure of the universe.

water hole (467) The interval of the radio spectrum between the 21-cm hydrogen radiation and the 18-cm OH radiation, likely wavelengths to use in the search for extraterrestrial life.

wavelength (75) The distance between successive peaks or troughs of a wave; usually represented by λ.

wavelength of maximum intensity (λ_{max}) (107) The wavelength at which a perfect radiator emits the maximum amount of energy; depends only on the object's temperature.

weak force (131) One of the four forces of nature; the weak force is responsible for some forms of radioactive decay.

west point (18) The point on the western horizon exactly halfway between the north point and the south point; exactly west.

white dwarf (272) The remains of a dying star that has collapsed to the size of Earth and is slowly cooling off; at the lower left of the H–R diagram.

Widmanstätten pattern (240) Bands in iron meteorites due to large crystals of nickel–iron alloys.

Wien's law (108) The mathematical relation between the temperature of an ideal radiator (a blackbody) and the wavelength at which it radiates most intensely.

WIMP (442) Weakly interacting massive particle, a hypothetical type of subatomic particle of which dark matter could be composed.

winter solstice (26) The point on the celestial sphere where the sun is farthest south; also, the time of year when the sun passes this point, about December 22, and winter begins in the Northern Hemisphere.

X-ray (76) Electromagnetic radiation with short wavelengths, high frequencies, and high photon energies, between gamma rays and ultraviolet radiation on the electromagnetic spectrum.

X-ray burster (353) An object that produces occasional X-ray flares. Thought to be caused by mass transfer in a closed binary star system.

Y dwarf (266) A spectral class of brown dwarf with temperature below 500 K.

Young Stellar Object (YSO) (302) A forming star in a late protostellar evolutionary stage that has lost its obscuring cocoon of gas but is still contracting toward the main sequence.

Zeeman effect (127) The splitting of spectral lines into multiple components when the atoms are in a magnetic field.

zenith (18) The point directly overhead on the sky.

zero-age main sequence (ZAMS) (311) The locus in the H–R diagram where stars first reach stability as hydrogen-burning stars.

zodiac (28) The band around the sky centered on the ecliptic within which the planets move.

Answers to Even-Numbered Problems

Chapter 1 **2.** 2160 mi; **4.** 1.1×10^8 km; **6.** about 1.2 seconds; **8.** 75,000 years; **10.** about 33 **Chapter 2** **2.** 2; **4.** 630; **6.** B is brighter than A by a factor of about 23; **8.** 66.6°; 113.4° **Chapter 3** **2.** (a) full; (b) first quarter; (c) waxing gibbous; (d) waxing crescent; **4.** 29.5 days later, about March 30; 27.3 days later, about March 28; **6.** 6840 arc seconds or about 1.9°; **8.** (a) The moon won't be full until about October 18; (b) The moon will no longer be near the node of its orbit; **10.** August 12, 2026 [July 10, 1972 + 3 × (6585 ⅓ days)]. Note that you must take into account the number of leap days in the interval to get the right answer. **Chapter 4** **2.** Retrograde motion: Mercury, Venus, Earth, Jupiter, Saturn, Uranus, and Neptune; never seen as crescents: Jupiter, Saturn, Uranus, and Neptune; **4.** $\sqrt{27} = 5.2$ years; **6.** about 6; no, the ratio in the Ptolemaic diagram is about 1.5. **Chapter 5** **2.** The force of gravity on the moon is about one-sixth of the force of gravity on Earth; **4.** 7350 m/s; **6.** 5060 s (1 hr 24 min); **8.** The cannonball would move in an elliptical orbit with Earth's center at one focus of the ellipse; **10.** 6320 s (1 hr 45 min) **Chapter 6** **2.** 3 m; **4.** Each Kecktelescope has a light-gathering power that is 1.6 million times greater than the human eye; **6.** No, his resolving power should have been about 5 arc seconds at best; **8.** 0.013 m (1.3 cm or about 0.5 in.); **10.** about 32 cm (from a distance of 400 km, a linear size of 0.7 m corresponds to an angular size of about 0.35 arc second.) **Chapter 7** **2.** about 150 nm; **4.** by a factor of 16; **6.** 250 nm; **8.** about 0.65 nm **Chapter 8** **2.** It will look $(2.06 \times 10^5)^2 = 4.2 \times 10^{10}$ times fainter; **4.** about 4×10^6 yr; **6.** overestimate by a factor of about 1.6; **8.** about 22 times; **10.** about 30 K; **12.** about 1.2×10^{-4} arc second **Chapter 9** **2.** 2.6 Mpc; **4.** 26 Mpc; **6.** about 1.6×10^8 y (160 million years); **8.** about 5.0×10^8 y (500 million years); **10.** 4.5×10^{41} kg (2.3×10^{11} solar masses) **Chapter 10** **2.** about 8×10^6 years; **4.** 0.024 pc; **6.** −28.5; **8.** about +16th mag; **10.** redshift $z = 0.16$ **Chapter 11** **2.** 57 km/s/Mpc; about 18 billion years; acceleration would decrease the age; **4.** 1.6×10^{-27} kg/m^3 (which is almost exactly one H atom per cubic meter); **6.** 76 km/s/Mpc **Chapter 12** **2.** It will look $(2.1 \times 10^5)^2$ = about 4.4×10^{10} times fainter, which is 26.6 magnitudes fainter; +22.6 mag; **4.** about 2.3 half-lives, or 3.0 billion years; **6.** large amounts of methane and water ices; **8.** about 1300 impacts per hour **Chapter 13** **2.** 720 km; **4.** 3.6 times brighter; **6.** 6×10^{14} J; **8.** 2.5×10^9 megatons; **10.** 4.3×10^9 kg **Chapter 14** **2.** 63 pc; absolute magnitude of 2.0; **4.** (a) B,15,000 K; (b) F, 6600 K; (c) M, 3000 K; (d) K, 4100 K; **6.** 10 pc; 33 ly; **8.** 3.87×10^{26} watts using 1.50×10^{13} m = 1 AU; Celestial Profile 1 gives 3.84×10^{26} watts; **10.** 1.3 solar masses; **12.** about 47 solar luminosities; about 220 solar luminosities; no, because mass and luminosity do not have a simple relationship for stars off the main sequence **Chapter 15** **2.** about 60,000 nm; **4.** 0.0001; **6.** about 1.5×10^6 years; **8.** The given values lead to an estimate of 930 years; the actual age is known to be about 960 years; **10.** 4.2 hours; 97 years **Chapter 16** **2.** about 24 km/s; **4.** 3.0 km/s; **6.** about 3.2 times the radius of the sun; **8.** There are four ^{1}H nuclei in the figure. They are used in a series of reactions to build up nuclei from ^{12}C to ^{15}N. In the last reaction (with the ^{15}N), a ^{12}C and a ^{4}He are produced. Because the ^{12}C can be used in the next cycle, the net result is one ^{4}He nucleus out for four ^{1}H nuclei in; **10.** 7.8×10^{33} kg (about 4000 solar masses) **Chapter 17** **2.** about 1×10^7 years for a 16-solar-mass star; about 6×10^5 years for a 50-solar-mass star; **4.** about 1×10^6 times less than present, or about 1.4×10^{-6} g/cm^3; **6.** about $(1/800)^3$ or 2×10^{-9}; **8.** about 3 pc; **10.** 183 seconds (about 3 minutes) early after 1 year **Chapter 18** **2.** about 1.8 ly; **4.** 29 Mpc; **6.** about 940 years ago (approximately the year 1070, close to the year 1054 in which the supernova was actually observed); **8.** about 2400 pc **Chapter 19** **2.** 7.1×10^{25} J/s or about 0.19 solar luminosity; **4.** 960 km/s (assuming the neutron star has 1.4 solar masses); **6.** about 11 arc seconds; **8.** about 490 seconds **Chapter 20** **2.** about 17 percent; **4.** about 8.2×10^7 yr (82 million yr); they have been subducted; **6.** 0.22 percent

Index

Bold page numbers indicate of key terms.

D

E

F

G

S

T